

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:10:02 ; Search time 49 Seconds  
(without alignments)  
5634.254 Million cell updates/sec

Title: US-10-005-480A-743  
Perfect score: 4804  
Sequence: 1 MESLTLTATEQPVKKNTLKK.....QPVSEILQLKYLPTFETTI 875

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriaph.\*  
17: sp\_archaeap.\*

Result No.	Score	Query Match %	Length	DB ID	Description
1	2540.5	52.9	565	11 Q99K07	Q99K07 mus musculus
2	2435	50.7	906	11 Q924C4	Q924C4 mus musculus
3	2191	45.6	874	13 Q7ZXN7	Q7zx7 xenopus lae
4	1360	28.3	251	4 Q9NCM9	Q9ncm9 homo sapien
5	1263	26.3	817	12 Q9J5H1	Q9j5h1 fowlpox vir
6	1263	26.3	817	12 Q9J5H1	Q9j5h1 fowlpox vir
7	972	20.2	415	11 Q8CAF0	Q8caf0 mus musculus
8	837	17.4	195	11 Q91VZ7	Q91vz7 mus musculus
9	835	17.4	152	4 Q9H515	Q9h515 homo sapien
10	755.5	15.7	479	10 Q9FS13	Q9fs13 spinacia ol
11	745	15.5	274	4 Q7Z3P5	Q7z3p5 homo sapien
12	743.5	15.5	496	10 Q9SU83	Q9su83 arabidopsis
13	737.5	15.4	479	10 Q4Z974	Q4z974 oryza sativ
14	713.5	14.9	457	10 Q84WJ3	Q84wj3 arabidopsis
15	712.5	14.8	457	10 Q9SU82	Q9su82 arabidopsis
16	680.5	14.2	461	10 Q9SU81	Q9su81 arabidopsis

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Match 52.9%; Score 2540.5; DB 11; Length 565;  
Best Local Similarity 81.1%; Pred. No. 3.1e-207;  
Matches 460; Conservative 53; Mismatches 51; Indels 3; Gaps 3;

QY 310 KAERPRFTYTFEPPDSSCHAGGVPVSARVIALQVVDHAFGMLMEGLKQRLHNCVNIL 369  
DB 1 KADRPSTYITTYVEBPSAGHSGSPVAGVIALQVVDHAFGMLMEGLKQRLHNCVNII 60

Result No.	Score	Query Match %	Length	DB ID	Description
17	677.5	14.1	151	11 Q9QYV2	Q9qyv2 rattus norv
18	671	14.0	257	11 Q88827	Q88827 rattus norv
19	644	13.4	477	11 Q921P7	Q921p7 mus musculus
20	644	13.4	477	11 Q95QG7	Q95qg7 mus musculus
21	637	13.3	485	3 Q943Z3	Q943z3 schizosacch
22	635	13.2	757	5 P90755	P90755 caenorhabdi
23	632.5	13.2	477	4 Q9UJA9	Q9uja9 homo sapien
24	630.5	13.1	429	2 Q69013	Q69013 zymomonas m
25	629.5	13.1	453	4 Q9Y6X5	Q9y6x5 homo sapien
26	623	13.0	829	5 P90754	P90754 caenorhabdi
27	607.5	12.6	456	11 Q8BTJ4	Q8btj4 mus musculus
28	604.5	12.6	453	11 Q8KIL3	Q8kil3 mus musculus
29	581.5	12.1	281	6 Q9TSB2	Q9tsb2 bos taurus
30	578.5	12.0	432	16 Q8PIS1	Q8pis1 xanthomonas
31	553	11.5	451	16 Q9A8S4	Q9a8s4 caulobacter
32	548.5	11.4	424	16 Q8P7F7	Q8p7f7 xanthomonas
33	518.5	10.8	440	4 Q96X57	Q96x57 homo sapien
34	510	10.6	440	11 Q8BGN3	Q8bgn3 mus musculus
35	510	10.6	458	4 Q8IU88	Q8iuh8 homo sapien
36	502.5	10.5	433	16 Q87A55	Q87a55 xyliella fas
37	499.5	10.4	433	16 Q9PAB9	Q9pab9 xyliella fas
38	356.5	7.4	234	10 Q94K08	Q94k08 arabidopsis
39	353.5	7.4	497	5 Q17540	Q17540 caenorhabdi
40	278	5.8	465	5 Q22129	Q22129 caenorhabdi
41	244	5.1	51	4 Q9NTT8	Q9ntt8 homo sapien
42	222.5	4.6	434	16 Q97LS7	Q97ls7 clostridium
43	221.5	4.6	431	16 Q8XKL9	Q8xkl9 clostridium
44	212	4.4	133	10 Q9SU80	Q9su80 arabidopsis
45	209	4.4	422	16 Q8V778	Q8v778 listeria mo

ALIGNMENTS

RESULT 1

Q99K07 PRELIMINARY; PRT; 565 AA.

AC Q99K07; 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Similar to ectonucleotide pyrophosphatase/phosphodiesterase 3 (Fragment)

DE ENFP3 OR A1876438.

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC005527; AAH05527.1; -.

DR MGI; MGI:2143702; Enpp3.

DR GO; GO:0004519; F:endonuclease activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0009117; F:nucleotide metabolism; IEA.

DR InterPro; IPR001604; Endonuclease.

DR InterPro; IPR002591; Phosphodiesterase.

DR Pfam; PF01223; Endonuclease; 1.

DR Pfam; PF01663; Phosphodiesterase; 1.

DR SMART; SM00477; NUC; 1.

FT NON\_TER

SQ SEQUENCE 565 AA; 63895 MW; 7179104FI8634A04 CRC64;

```

QY 370 LADHGMDDQTYCNKWEYMTDFFPRINFFYMEGPAIRAHNPHDPFFSNSEIIVRNLS 429
Db 61 LADHGMDDQTYCNKWEYMTDFFPKIN-FYMQGPAIRAHNPHDPFFSNSEIIVRNLS 119
QY 430 RKPDDHFKPYLTPDLPRKLYAHKVRIDKVLHVFQOQWLAVERSKNTNCGGNGHYNEF 489
Db 120 RKPDDHFKPYLTPDLPRKLYAHKVRIDKVLHVFQOQWLAVERSKNTNCGGNGHYNEF 179
QY 490 RSMELFLAHGSPFKETVEPENIEVYNMCDLLRIOPAFNNGTHGSLAHLLKVPVE 549
Db 180 KSMELFLAHGSPFKETVEPENIEVYNMCDLLRIOPAFNNGTHGSLAHLLKVPVE 239
QY 550 PSBAEVSFSGVGFANPLFTESLDCFCPLHQLNLTQ-BOVNQMLNLTQEEITATVKVL 608
Db 240 PSBAEVSFSGVGFANPLFTESLDCFCPLHQLNLTQ-BOVNQMLNLTQEEITATVKVL 299
QY 609 PGRPRVLONVDHCLLYHREYVSGFGKAMRPMHSSYTFVPGDTSPLPTVPDCLRAD 668
Db 300 PGRPRVLONVDHCLLYHREYVSGFGKAMRPMHSSYTFVPGDTSPLPTVPDCLRAD 359
QY 669 VRVPSESOKCSFYLADKNIITHGFLYPPASNETSDQDALITSNLVPMYEFKQWMDYF 728
Db 360 VRVPSESOKCSFYLADKNIITHGFLYPPASNETSDQDALITSNLVPMYEFKQWMDYF 418
QY 729 HSVLLIKHATERGNVNVSGPIFDYNYOGHPADPEITKHLANTDVPITPHYFVVLTSCK 788
Db 419 HEVLLIKHATERGNVNVSGPIFDYNYOGHPADPEITKHLANTDVPITPHYFVVLTSCK 478
QY 789 NKSHTPENCPCGWLDPFTIIPHRPTNVESCPGKPEALWVERFTAHARVDRDVELLTL 848
Db 479 DQTHPDCSCGWLDPFTIIPHRPTNVESCPGKPEALWVERFTAHARVDRDVELLTL 538
QY 849 DFYQKQVPSBILQKTYLPTFTT 875
Db 539 DFYQKQVPSBILQKTYLPTFTT 565

RESULT 2
Q924C4 PRELIMINARY; PRT; 906 AA.
AC Q924C4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 1 allotype b
DE (EC 3.1.4.1).
GN ENPPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
SEQUENCE FROM N.A.
RA Banakh I., Sai A., Dubljevic V., Grobten B., Slegers H., Goding J.W.;
RT "structural basis of allotypes of ecto-nucleotide
RT pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein
RT PC-1) in the mouse and rat";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF339910; AAK84174.1; -.
DR MGD; MGI:97370; Enppi.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiester.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01223; Endonuclease; 1.
DR Pfam; PF01663; Phosphodiester; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR SMART; SM00477; NUC; 1.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW Hydrolase.

```

```

SQ SEQUENCE 906 AA; 103125 MW; 1FA071071B4B9C50 CRC64;
Query Match 50.78; Score 2435; DB 11; Length 906;
Best Local Similarity 51.04; Pred. No. 5.8e-198;
Matches 447; Conservative 152; Mismatches 250; Indels 28; Gaps 9;
QY 10 EQPVKK-----NTLKKYKIACIVLLALVIMSLGLGLGLGRKLEKQSGCKRCKCF 59
Db 39 EEPLEKAEARPAKDPNT---YKVLVLVSVCLVTILGICIFGLKPKSCAKVKSKGRCF 95
QY 60 DASPRGLENCRCVACHKRGDCWDEPDTCTVETRIWMKNKPCGTRLEASLCSSDDC 119
Db 96 ERTF---SNCRCDAAVCVSLGNCCLDFQETCTVBETHIWTCKNPKCGEKLRSFVCSADDC 152
QY 120 LQKDCDCADYKVCOCGETSWELENCDTAQOSQCEGFDLPPVILFMSDCGFRAEYLTWDT 179
Db 153 KTHNDCCINYSVCOQDKSWVEETCESIDTPECPAEFESPTLFLDLDFRABYLHTWG 212
QY 180 LMPNINKLTCGTHSKYMRAMYPKTTPNNHYTITGLYPESHGIIIDNNYDVNMLNKFSL 239
Db 213 LLPVISKLXKCGTYTKMRPMYPTKTFPNHYSITVGLYPESHGIIIDNNYDVNMLNKFSL 272
QY 240 SSKEONNPAMHOGPWLTAQGLKAAATYFPGSEVAINSGSPSYMPVNGSVPEERI 299
Db 273 KSKEFNPLWKQPIWVTANHOEVKSGTYFPGSDVEIDGILPDIYKYNVSGVPEERI 332
QY 300 STLLKWLDPKAEPRPYTYMFEPPDSSGHAGPVSARVICALQVVDHAFGLMEGLKOR 359
Db 333 LAVLEWLQPSHERPHYTYLLEPPDSSGHSGHSPVSSEVIKALQKVDRLVGLMDGLKOL 392
QY 360 NLHNCVNIILLADHGMDDQTYCNKWEYMTDFFPRINFFYMEGPAIRAHNPHDPFFSN 419
Db 393 GLDKCNLNLISDHGMEQSGCKRYVYLNKYLGVNWNKVYVGYGAARLRTDVPETYSN 452
QY 420 SEBIVRNLSCKPQDQHFKPYLTPDLPRKLYAHKVRIDKVLHVFQOQWLAVERSKNTNC 478
Db 453 YEALAKNLSCEPNQHFRPYLKPFLPRKLYAHKVRIDKVLHVFQOQWLAVERSKNTNC 512
QY 479 GGNHGYNNEFRMEAI FLAHGSPFKETVEPENIEVYNMCDLLRIOPAFNNGTHGS 538
Db 513 GSGFHGSDNLFNMQALFICYGPAFKHGAEDVSFENIEVYNMCDLLRIOPAFNNGTHGS 572
QY 539 LNHLLKVPFVPSHAEVSKFSVCGFANPLFTESLDCFCPLHQLNLTQEEITATVKVL 598
Db 573 LNHLLKVPFVPSHAEVSKFSVCGFANPLFTESLDCFCPLHQLNLTQEEITATVKVL 628
QY 599 EITATVKVNLPPGRPRVLONVDHCLLYHREYVSGFGKAMRPMHSSYTFVPGDTSPLP 658
Db 629 DVDDIYHMTVPYGRPRILLKQHRVCLLQOQQLTQOQQLTQOQQLTQOQQLTQOQQLT 686
QY 659 PTVPDCLRADVRVPSESOKCSFYLADKNIITHGFLYPPASNETSDQDALITSNLVPM 717
Db 687 DPFNSCLYQDLRIPLSPFVHKSKSYKSNKLSYSGFTLTPPRLNRSVNHVSEALLTNIVPM 746
QY 718 YSEFRKMWYFHSVLLIKHATERGNVNVSGPIFDYNYOGHPADPEITKHLANTDVPITPHY 774
Db 747 YQSFQVIMHVLHDTLLQRYAHERNGINVGVPFDFDYDGYDGYDGYDGYDGYDGYDGYD 806
QY 775 PIPTHYFVVLTSCKNKSHTPENCPCGWLDPFTIIPHRPTNVESCPGKPEALWVERFTAH 834
Db 807 LIPTHTFFVLTCKQLSETPLECSA-LESSAVILPHRPDNIESTCHGKRESSWVEELLTL 865
QY 835 HIARVDRDVELLTLGDFYQKQVPSBILQKTYLPTFTT 871
Db 866 HRAVTDVELLTLGDFYQKQVPSBILQKTYLPTFTT 902

RESULT 3
Q7ZXN7 PRELIMINARY; PRT; 874 AA.
AC Q7ZXN7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
KW Hydrolase.

```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044675; AAH44675.1;
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0009117; F:nucleotide metabolism; IEA.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiesterase.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01223; Endonuclease; 1.
DR Pfam; PF01663; Phosphodiester; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR SMART; SM00477; NUC; 1.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW Hypothetical protein.
SQ SEQUENCE 874 AA; 99650 MW; 7EA5B4EC1D9FCE46 CRC64;

Query Match 45.6%; Score 2191; DB 13; Length 874;
Best Local Similarity 45.1%; Pred. No. 3.2e-177;
Matches 409; Conservative 161; Mismatches 255; Indels 82; Gaps 14;

QY 15 KNTLKKYKACVILALLVIMSLGLGLGLRK-----LEKQSCRKCC 58
DB 4 KNGFSFKHVISVLTFAIGINVLGFTANFRKSEBDEGVASVLSDSPIWSSGSKERC 63
QY 59 FDSAPRGLNCRDCAKDRGCCWDFEDTCVSTRIMWKNFRGCTRLASLSCSDD 118
DB 64 FELIEAABACRDNLCNSCCEDFDEHCLKTGRGECTKDRGCTRNEENACHCED 123
QY 119 CLKQDCCADYKSVCGQTSWLENCDTAQOCCQCEGPDLPVILFSDMGFRAEVLTYWD 178
DB 124 CLAKGDCCTNYQVCKGTHWADDDCEEMKHCEPCAGFVRPLLIIFSDVGRASYMKKH 183
QY 179 TLMENINKLCTGHSKYNRMVPTKTPNHVTITGLYPBSHGIDNNVDVNLKNFS 238
DB 184 KYNPNIDKRSCTGTHSPYWRPVPTKTPNLTALGLYPESHGIVGSMVDPVFDANFS 243
QY 239 LSSKEQNPANWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSPYNGSVPFBER 298
DB 244 LRSREKFNHRWGGQPIWITASKQGLKAATFPWP---VAISQ-----ERR 285
QY 299 ISTLLKWLDPKAEPRFTYMTYFEBPDSGSHAGGVSARVICALQVDPHAFQMLMEGLKQ 358
DB 286 IFTVLQWLHPDNERFYVVALYSEQDQAGHKYQFPQPELAQLKNDKIVQQLMDGLKQ 345
QY 359 RVLHNCVNTILLADHGMQDTCNKMEDYFPRINFFYMWEGPAPRAHNR--IPHPDF 416
DB 346 MXLHRCVNVIFVGDHGWBEATCERTEFLNSYLSNVDDFALLPGSIGRMSRPANKHD-- 403
QY 417 SFNSEIERNLSCKRPDQHFQKPYLTPDLPKRLHYAKNVRIDKVHLFVDOQWLAVRS---- 472
DB 404 ---PKAVANLTKCKPDQHFQKPYLTKOHLKRLHYAFNRRIEDIHLLVDKRWKHAVKPMVD 460
QY 473 -KSNTWCG- GNGHYNNEFRSMEAFIPLAHGPFSEKTEVEPENTEVNLMCDLLRIOPA 530
DB 461 YKQKQKACQGDHGYDNKITSQCTVFLGHGPFQFKYKTVPPENTELYNVACDVGLKPA 520
QY 531 PNNGTHGSLNHLKVPFYPSPSHAEVSKFSVCGFANPLP-----TESLDCFCPHLQNS 582
DB 521 SNNNGTHGSLNHLRVAKYKPAIPDEYSK-----PLPIVTSPTVNBELGSC---DD 569

583 STQLQVQNQL-----NLTOBEITATVK-----VNLPGRPVRLQKNDVHCLLYHR 628
570 KNKAEELNKLKLGTDVAVEELSNEIKELTSRNTDKNLLYGRPAVLKY-KTYSVLHHS 628
629 EYVSGFGKAMRMPWSSYTPQIGDTSPLPPTVPDCLRADRVRPVPSSEKSCSYLADKNI 688
629 DFESGFSESLMPLATSYITISKQADYSGIPEHLSNCRVLDPRISPGNSQSCSAYKADQM 688
689 THGLYPPASNRSTDSQYDALITSNLVPMYEBFRKWDYFHSVLLIKHATERGNVVS 748
689 SYGLFPPQLSSADSADSKYDPLTINPIYPAFKKIWNVYFORVLVRFATERGNVVS 748
749 PIPDYNVDGHFADPEITKHLANTDPIPTHYFWLITSCNKSHTPENCPCGMDLVLPFII 808
749 PIFDYDYGVDYDNDKI-KMFVDGSPVPTHYIITSCWDFNQAVDNCDCGLSVSVFVI 807
809 PHRTNVESCEPGKPEALWYERFTHAHARVDRVELLTGLDFYQDKVQPVSEILQIKTYL 868
808 PHRPDNDESCNSEESKWEKVDLLKXHTTIRDIETLLTGLDFYRKTNRSYTEILSLKTYL 867
869 PTPETTI 875
868 HTVESEI 874

RESULT 4
Q9NQM9 PRELIMINARY; PRT; 251 AA.
AC Q9NQM9, 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ1005H11.3 (Phosphodiesterase 1/nucleotide pyrophosphatase 3)
DE (Fragment).
GN PNP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lovell J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI35904; CAB99214.1;
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0009117; F:nucleotide metabolism; IEA.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiesterase.
DR Pfam; PF01663; Phosphodiester; 1.
DR SMART; SM00477; NUC; 1.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28265 MW; C726FD4719D4880C CRC64;

Query Match 28.3%; Score 1360; DB 4; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 SKSNTWCGNGHYNNEFRSMEAFIPLAHGPFSEKTEVEPENTEVNLMCDLLRIOPAP 531
DB 1 SKSNTWCGNGHYNNEFRSMEAFIPLAHGPFSEKTEVEPENTEVNLMCDLLRIOPAP 60
QY 532 NNGTHGSLNHLKVPFYPSPSHAEVSKFSVCGFANPLPTESLDCFCPHLQNSQLEQVQ 591
DB 61 NNGTHGSLNHLKVPFYPSPSHAEVSKFSVCGFANPLPTESLDCFCPHLQNSQLEQVQ 120
QY 592 MLNLTQBEITATVKVNLPGRPVRLQKNDVHCLLYHREYVSGFGKAMRMPWSSYTPQL 651
DB 121 MLNLTQBEITATVKVNLPGRPVRLQKNDVHCLLYHREYVSGFGKAMRMPWSSYTPQL 180
```

```
QY 652 GDTSPPLPTVPDCLRADVRPSESOKCSFYLLADKNITHGFLYPPASNRTSDSOYDALIT 711
|||||
Db 181 GDTSPPLPTVPDCLRADVRPSESOKCSFYLLADKNITHGFLYPPASNRTSDSOYDALIT 240
|||||
QY 712 SNLVPWYEEFR 722
|||||
Db 241 SNLVPWYEEFR 251
|||||

RESULT 5
Q9J5H1
ID Q9J5H1 PRELIMINARY; PRT; 817 AA.
AC Q9J5H1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF FPV030 alkaline phosphodiesterase.
GN FPV030.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
RL [2]
RP SEQUENCE FROM N.A.
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF198100; AAF44374.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiester.
DR Pfam; PF01663; Phosphodiester; 1.
DR SMART; SM00477; NUC; 1.
DR NCBI_TaxID=10261;
SQ SEQUENCE 817 AA; 94038 MW; 8404FD00641DA022 CRC64;

Query Match 26.3%; Score 1263; DB 12; Length 817;
Best Local Similarity 36.9%; Pred. No. 2.6e-98;
Matches 271; Conservative 145; Mismatches 279; Indels 40; Gaps 18;

QY 149 QSQCEGFDLPPLVFLSMGDFRAEYLYTWDTLMPNINKLTCGHSKYMRYPTKTFPN 208
|||||
Db 108 QVTCPEFERPPLILIAMNGFRDYLNKWEKIPTIKDLMHGVTAP-MRPVYPTNTFPN 166
|||||
QY 209 HYTVITGLYPESHGIIDNNYDNLNKNFSLSSKEQNNPAMWGHQPMWLTANYOGLKAAT 268
|||||
Db 167 LYSIVTGLYPISHGIDNEFDIRGDTIEFTIASEEVEWFGGEPIWTTIMKNGFKSAT 226
|||||

QY 269 YFWPGSEVAINGSFSPYIYNGSVPPFERISTILKWLDPKAPRPFTYTFYFSEPOSSG 328
|||||
Db 227 FFWPGSDKVPKRPFTYRSYNSKVPYERINTVLRLKMDTGYPFYALYLEPSSG 286
|||||

QY 329 HAGGVSARVIALQVDFHAFGLMEGLKQNLNHCNVIILADHGMDOQTCNKNRYWTD 388
|||||
Db 287 YEYGTDDERVGKALEKVDKATALLMGLKDLQGLICANLILSDHGMNSVDPKIVNLKD 346
|||||

QY 389 YPRNFFVMYEGAPAPRAHNIPIHDFFSFNSSEIVRLNLSCKPQDQHPKPYLTLDLPRL 448
|||||
Db 347 YITN-NDVVIKGATPVIKPQNLH-IRLFYDGLIISTSCWDDQDFIVYSRRLPRL 404
|||||

QY 449 HYAKNVRIDKVLHFDVQOQLAVRSKN-TNCGGNHGYNNFRSMEAIFLAHGSFKEKT 507
|||||
Db 405 HYGSGFRTEILGVYLEEGQSTDENGLKRSRGGFGHSDNSFQDMTAVFLGYGPAFLDDV 464
|||||

QY 508 EVERPENIEVNLMDLLRIQAPNNGTHGSLNHLKVPFPEPSSHAEVSKFSCGFPANP 567
|||||
```

```
Db 465 RVPIFDNIELYNNMCEILGINPANNNGTVGSLNNILNRGK--THVSSLDIITI----- 516
|||||
QY 568 LPTESLDC-----FCHQLNSTOLEQVQNLNLTQBEITATVK-----VNLPRGRPRVLQ 617
|||||
Db 517 --ES-ECDRHAYVGDHLKGCT-CKNIDRFSSKGNKEDSSRTRSSSYIYNLPFGKPAVLL 571
|||||
QY 618 KNVDHCLLYHREYVSGFGKAMRPMKSSYTVP-QLGDTSPPLPTVPDCLRADVRPSES 676
|||||
Db 572 NRHHCHIKNDNVYTAYSKVRNPLATSFSDITYNSTIYNTK---CYLQDMRV-MYK 627
|||||
QY 677 QKCSFYLLADKNITHGFLYPPASNRTSDSOYDALITNSLVPMBZEEFRMDYFHSVLLIKH 736
|||||
Db 628 EPCRYSTQKDVYGYLYPA---RAITD--FQSLLETNTPVMYRNFKKIWEVFMSSILIEY 682
|||||
QY 737 ATERNGVNVSGPIFDYNYDGHGDPADDEITKHLA-NTDVPITHYFWLITSCQKSHTEPE 795
|||||
Db 683 VQKHVVNVNVGVPFDSNSNGIRDSWLISMSSGYNNKVIIPSDYFVILTYCKDKDSSLN 742
|||||
QY 796 NCPGWLDDVLPFIIPHRPTNV-ESCP-EGKPEALMVEERTAHIAHVRDVELTGLDFYQD 853
|||||
Db 743 DCYSNKTESFVVPNSDTTYNESCSENITSVSVYRKIFSLHRVRIKDIETVTSMSFYRN 802
|||||
QY 854 KVQPVSEILQLKTYL 868
|||||
Db 803 VYKTSNVAYLKTIM 817
|||||

RESULT 6
O90761
ID O90761 PRELIMINARY; PRT; 817 AA.
AC O90761
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alkaline phosphodiesterase I (EC 3.1.4.1).
GN PCL.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RA Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;
RT "Fowlpox virus encodes non-essential homologs of cellular alpha-SNAP,
PC-1 and an orphan human homolog of a secreted nematode protein.";
RL J. Virol. 72:6742-6751(1998).
RL [2]
DR EMBL; AJ006408; CAA07014.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004528; F:phosphodiesterase I activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiester.
DR Pfam; PF01663; Phosphodiester; 1.
DR SMART; SM00477; NUC; 1.
DR NCBI_TaxID=10261;
RX SEQUENCE 817 AA; 94004 MW; CB68DA0508CC568E CRC64;

Query Match 26.3%; Score 1263; DB 12; Length 817;
Best Local Similarity 36.9%; Pred. No. 2.6e-98;
Matches 271; Conservative 145; Mismatches 279; Indels 40; Gaps 18;

QY 149 QSQCEGFDLPPLVFLSMGDFRAEYLYTWDTLMPNINKLTCGHSKYMRYPTKTFPN 208
|||||
Db 108 QVTCPEFERPPLILIAMNGFRDYLNKWEKIPTIKDLMHGVTAP-MRPVYPTNTFPN 166
|||||
QY 209 HYTVITGLYPESHGIIDNNYDNLNKNFSLSSKEQNNPAMWGHQPMWLTANYOGLKAAT 268
|||||
Db 167 LYSIVTGLYPISHGIDNEFDIRGDTIEFTIASEEVEWFGGEPIWTTIMKNGFKSAT 226
|||||
```



269 YFWPGSEVAINSPFISYMPYNGSVPPFEERISTLLKWLDPKAERPRFYTMVFEPPDSG 328  
227 FFWPGSKVPPKRPMTYRSYKNSVPYEBERINTVLRWLKMDTGYPYFYALYLEBPGSSG 286  
329 HAGGPVSARVIALOVVDEAFGLMEGLKQRLNHCNVAILLADHGMDOQTCYKNEYMTD 388  
287 YEGTDDERVGKALEKVDKAILMKGLKDLIGCANLILVSDHGMNVDPFKITVNLKD 346  
389 YPRINFFMYGAPAPRIAHNIPIHDFSFNSEEEIVRLNSCRKPQHFQKPYLTDPKRL 448  
347 YITN-NDVVIKPGATPVKPNLH-IRLFDVIGIISSTSCVMDQPPFIVSVRSLPKRL 404  
449 HYAKNVRIDKVLFDVDOQWLAVRKSN-TNCGGNGHGVNNEPRSMEAIFLAHGPSFKKT 507  
405 HYSGFRTEILGYLEEKGQSDDENGLKHSRGGFHGSDNSFQDNTAVFLGYGPAFLDDV 464  
508 EYEPFENIEVYNLMDCLLRIQAPNNGTHGSLNHLKVPFYPBPSHAEVSKFSVCGFANP 567  
465 RVPIFCNIELYNMCEILGINPANNNGTVGSLNHLNRSRY--THVSLDIITI----- 516  
569 LPTESLDC-----FCHLQNSTQLBOVQMMLNLTQEEITATVK-----VNLPGPRVLQ 617  
517 ---ES-ECDRHAYVGDHLKGCT-CXNIDRFSSKGNKEDSSRTRSSSYIYNLPFGKPAVLL 571  
618 KNVDCLLYHREYVSGFGKAMRWKSSYTPV-QLGDTSPLPPTVPDCLRADVRPPSES 676  
572 NRHHCILKNDYVTAISKVNLPLWTFSIDITNSTNINYNK---CYLQDMRV-MYK 627  
677 QKCSFYLDKNTHTGFLYPPASNRSDSQYDALITSNLPVMEYEPKQWDFHSLVILIKH 736  
628 EPCRYVSTQKVTGYLYPA---RATD--FQSLLETNTVMYRNFKKIWEVFMSSILLEY 682  
737 ATERNGVNVSGPIEDYNDVGHDFDAPDEITXHLA-NTDVPITPHYFVVLTSCKNKSHTPE 795  
683 VQKHVVNVVGVFDSNSGRDSCWLISMSSGNNKVPYSPDVFILTYCKDKDSSLN 742  
796 NCGPGLMDLVPFPIPRPTNV-BSCP-EGKPEALWVEERFTHAIAVRDVELLITGLDFYQD 853  
743 DCYSNIKTESFVVPNSDTYNESCSRENITSVYRKIFSLHVRVRIKIEIVTSNFSYRN 802  
854 KVQPVSEILQLKTYL 868  
803 VYKTSNVAYLKTYM 817

RESULT 7  
Q8CAFO PRELIMINARY; PRT; 415 AA.  
AC Q8CAFO;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Ectonucleotide pyrophosphatase/phosphodiesterase 2.  
GN ENP2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Hypothalamus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573 (2002).  
DR EMBL; AK038940; BAC30174.1; -.  
DR MGD; MGI:1321390; Enpp2.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0009117; P:nucleotide metabolism; IEA.  
DR InterPro; IPR002591; Phosphodiesterase.  
DR InterPro; IPR001212; Somatomedin\_B.

DR Pfam; PF01663; Phosphodiesterase 1.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR PRINTS; PR00022; SOMATOMEDIN\_B.  
DR SMART; SM00201; SO; 2.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
SQ SEQUENCE 415 AA; 47056 MW; BCA9A7742C6A3B60 CRC64;  
Query Match 20.2%; Score 972; DB 11; Length 415;  
Best Local Similarity 41.0%; Pred. No. 5.7e-74;  
Matches 172; Conservative 72; Mismatches 124; Indels 52; Gaps 4;  
QY 32 LVMSIGLGLGLG-----LRKLE-----KQSGCRKCKCFDASFGLENCR 70  
DB 15 LFTFAIGNLCLGFTASRIKRAEWDGPPTVLSDSPWNTSGCKGRCFELQEVGPPDCR 74  
QY 71 CDVACKDRGCCDFEDTCVETRIWCMKFRGCTREASLCSGSDDCQKQKCCADYK 130  
DB 75 CDNLCKSYSSCCHDFDLCLKTARGWECTKDRGGEVRNEENACHCEDCLSRGDCCTNYQ 134  
QY 131 SVCGETSWLEENCDAQAQSCPEGFDLPVILFSDMGFRAEYLYTWTLMPNINKLKC 190  
DB 135 VCKGESHWDVDDCEIEIRVPECAPGVRRPLLIIFSVDFGFRASYMKKSKVMPNIEKLRS 194  
QY 191 GIHSKYNRMYPTKTFENHYTIVTGLYPESHGIIIDNNYDVNLNKFSLSSKQNPAAW 250  
DB 195 GTHAPYMRPVPTKTFENHYTIVTGLYPESHGIIIDNNYDVNLNKFSLSSKQNPAAW 254  
QY 251 HGQPMMLTAMYQGLKAATYFPGSEVAINGSPFISYMPYNGSVPPFEERISTLLKWLDPK 310  
DB 255 GGQPLWITATKQGVRACTFEW-----SVSIPHERRILUTILQWLSLPD 296  
QY 311 AERPRFTMYFEPPDSGSHAGGVPVSAVIALOVVDEAFGLMEGLKQRLNHCNVAILL 370  
DB 297 NERPSVAFYSEQDFSGHYGPGPEMTNPLEIDKTVQQLMDGLKQLKHLRCVNVIFV 356  
QY 371 ADHGMDOQTCYKNEYMTDYPPIRNFYFMYEGPAPRIAHNIPIHDFSFNSEEEIVRLNSCR 430  
DB 357 GDHGMEDVTCRTFEFLSNLYLTNVDDITLVPGTLGRIR-----PKIPNNLKCK 403  
RESULT 8  
Q91VZ7 PRELIMINARY; PRT; 195 AA.  
AC Q91VZ7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Similar to alkaline phosphodiesterase.  
GN ENP3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Breast tumor;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC006944; AA06944.1; -.  
DR MGD; MGI:2143702; Enpp3.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0009117; P:nucleotide metabolism; IEA.  
DR InterPro; IPR002591; Phosphodiesterase.  
DR InterPro; IPR001212; Somatomedin\_B.  
DR Pfam; PF01663; Phosphodiesterase 1.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR SMART; SM00201; SO; 2.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
SQ SEQUENCE 195 AA; 21541 MW; D64F2C37EFE4A65 CRC64;  
Query Match 17.4%; Score 837; DB 11; Length 195;  
Best Local Similarity 75.9%; Pred. No. 5.9e-63;  
Matches 145; Conservative 21; Mismatches 25; Indels 0; Gaps 0;



```

Query Match          15.5%; Score 746; DB 4; Length 274;
Best Local Similarity 54.5%; Pred. No. 5.3e-55;
Matches 140; Conservative 37; Mismatches 78; Indels 2; Gaps 2;

QY 322 EEPDSSGAGGVSARVIALQVVDHAFGMLMEGLKQRLNHCNVIILLADHGMDQTYCN 381
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 EEPDSSGSHSGVSSEVIKALQVDMVGMMDGLKELNLRCLNLIILSDHGMEOGSK 62
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 382 KVEYMTDYPPRINFFVMEGAPAPRIRAHNI PHDFFSFENSEEIVRNLSCKRQDQHKPYLT 441
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 KYIYLNKYLGVDKNIKVIYGPAPRIRAHNI PHDFFSFENSEEIVRNLSCKRQDQHKPYLT 122
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 442 PDLKELHYAKNVRIDKVLHFDVQW-LAVRSKSNTCGGNGHNGYNNFRSMEAI FLA 500
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 HFLPKLHFAKSDRIEPLTFYLDQWQLALNPSERKYCGSGPHGSDNVFSNQALFVGYG 182
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 501 PSFKETETEVEPENIEVNLMDLLRIQAPNNGTHGSLNHLKVPFFPEPSHAEVSKFS 560
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 PGFKHGEADTFENIEVNLMDLLRIQAPNNGTHGSLNHLKVPFFPEPSHAEVSKFS 242
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 561 VCGFANPLTESLDCFC 577
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 QCFPTFN-PRDLGSC 258
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
Q9SU83 PRELIMINARY; PRT; 496 AA.
AC Q9SU83;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nucleotide pyrophosphatase-like protein (EC 3.6.1.9).
GN T16L4.190 OR A74G29680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA Mayer K.P.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.P.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EMBL; AL079344; CAB45328.1; -.
DR EMBL; AL161575; CAB79726.1; -.
DR FIC; T02931; T02931.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004551; P:nucleotide diphosphatase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiester.
DR Pfam; PF01663; Phosphodiester; 1.
KW Hydrolase.
SQ SEQUENCE 496 AA; 54678 MW; 8DC2B4346121D732 CRC64;

Query Match          15.5%; Score 743.5; DB 10; Length 496;
Best Local Similarity 41.0%; Pred. No. 2e-54;
Matches 159; Conservative 71; Mismatches 141; Indels 17; Gaps 9;

QY 157 DLPPIVFLSMDGPRASYLYTWDTLMPNINKLTCGLHSKY-MRAMYPTKTFNHYITVG 215
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 105 DKPWLLISSDGRF--FGYQKTKLPISIRHLIANGTEAGTGLIPVFTLTLPNHYISVTG 162
QY 216 IYPESHGIIINMYDVNINLKNFSLSSKEQNPNAMWEGOPMWTAMYOGLKAATYFWPGSE 275
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 LYFAYHGIINHHFVDBETGNVFTMASHE---PEWMLGEPJWETVWQGLKAATYFWPGSE 219
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 276 VAINGSF---PSIYMPYNGVSPFEERISTLLKWLDPKAEPRFRFYTYMFEPPSSGHAGG 332
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 220 VH-KGSWNCQGLCQNVGSPEDDRVDTILSYFDLPSEIIPSFMTLYFPDQHGQVHG 278
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 333 PVSARVIALQVVDHAFGMLMEGLKQRLNHCNVIILLADHGMDQTYCNKMEYMTDYFPR 392
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 PDPQITEAVYNTDRILGRIDGLGKRGIFEDVTIMVGDHGMVGTCDKLVVLDLAPW 338
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 393 INF---PYMYEGAPAPRIRAHNI PHDFFSFENSEEIVRNLSCKRQDQHKPYLT 447
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 IKPSSNVQYITPLAIQPPS-GHDAADIVA-KINGLSSGKVENGLKYLKEDLP 396
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 448 LHYAKNVRIDKVLHFDVQW-LAVRSKSNTCGGNGHNGYNNFRSMEAI FLA 507
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 397 LHYVDSDRIPPIIICLVDEGFKVEKSKAKCEGAGHYDNAFFSMRTIFIGHGFMFSKGR 456
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 508 EVEFFENIEVNLMDLLRIQAPNNGT 535
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 457 KVPSEFNVQIYNVISSILGKAAPNNGS 484
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
Q42974 PRELIMINARY; PRT; 479 AA.
AC Q42974;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nucleotide pyrophosphatase precursor (EC 3.6.1.9) (OSE4) (Nucleotide
DE pyrophosphatase homolog).
GN OSE4 OR B0419B01.2;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]_
RP SEQUENCE FROM N.A.
RA Hsing Y.-I.C., Hsu T.-F., Chen Z.-Y., Chung M.-C., Tsou C.-H.,
RA Hsieh J.-S.;
RT "Characterization of a rice early embryogenesis specific gene OSE4.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. IR64;
RA Hsing Y.-I.C., Hsu T.-F., Chen Z.-Y., Chung M.-C., Tsou C.-H.,
RA Hsieh J.-S.;
RT "Characterization of a rice early embryogenesis specific gene OSE4.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0419B01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
DR EMBL; U25430; AAA67067.1; -.
DR EMBL; AF245483; AA65458.1; -.
DR EMBL; AF003244; BAB56086.1; -.
DR FIC; T03293; T03293.
DR Gramene; Q42974; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004551; P:nucleotide diphosphatase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiester.
DR Pfam; PF01663; Phosphodiester; 1.
KW Hydrolase; Signal.

```



Query Match		14.8%;	Score 712.5;	DB 10;	Length 457;
Best Local Similarity		38.5%;	Pred. No. 7.8e-52;		
Matches 166;		Conservative 76;	Mismatches 148;	Indels 41;	Gaps 12;
QY	148	QSQCPGEGDLPVILFSDGPRABLYLTWDTLMNINKLTCGIHSKY-NRAMYPKTP	206		
Db	44	RRQPPKLNKEVLLISCDFRGYQPKTET--ENIDLLISRGTEAKTGLIPVPTMTF	101		
QY	207	PNHYTIVTGLYPESHGIIDNNMYDNLKNFSLSSKEQNNPAWHGQPMWLTAMYQGLKA	266		
Db	102	PNHYIATGLYPASGIIIMNKETDPVSGELFY-----RNINPKMWLGEPLWTVAVNQGLMA	157		
QY	267	ATYFWPGSEVAINGSF--PSIY--MPYNGSVPFEEIRSTLLKWLDPKAERPRFVTMYFE	322		
Db	158	ATYFWPGADVH-KGSWNCCKGCKAPXNVSVPLEERVDTILNYFDLPEREIPDFMAYFD	216		
QY	323	EPDSSGHAGGPVSARVICALQVVDHAFGLMEGLKQRLNHCNVIILLADHGMDOITYCNK	382		
Db	217	EPDIOGHEYGPDPRVTEAVSKVDXMI GRIIMGLEKRVFSDVAVILLGDHGMVTCNCKK	276		
QY	383	MEYMTDYFRINFFYMEGPAPRIRAHNIPHDFFSFN-----SEEIVRN---	426		
Db	277	VIIYIDDLADWIKI-----PADWIODYS---PVLAMNPRWGKDVKNPGQKNAELVRKWE	327		
QY	427	-LSCKR--PDQHKPYLTPDLPKRLHYAKNVRIDKXHLFVDOQWLAVRSKNTNCGGNH	483		
Db	328	ALSSGKVANGEFLQVLYLKENLPQRLHYSDSSRIPIIGMWGEGLMVQKQRTYVQECSTH	387		
QY	484	GYNNEFRSMEATFLAHGPSFKEKTEVEPPENIEVYNLMDLLRIOPAPNNGTHGSLNHL	543		
Db	388	GYNMFFSNRSIFVGYGPRFRGICKVPSFENVQVTVNAVAEILGLRPAPNNGSSLFTRSLL	447		
QY	544	KVPFVEPSSHAE	554		
Db	448	-LPFGETSQVE	457		

Search completed: July 6, 2004, 13:19:49  
Job time : 52 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 12:47:02 ; Search time 18 Seconds  
(without alignments)  
2531.188 Million cell updates/sec

Title: US-10-005-480A-743

Perfect score: 4804

Sequence: 1 MESTLTATREQVKNTLKK.....QPVSEILQKLYLPTFFETI 875

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4804	100.0	875	1 NPP3_HUMAN	O14638 h ectonucle
2	3982	82.9	875	1 NPP3_RAT	P97675 r ectonucle
3	2498.5	52.0	925	1 NPP1_HUMAN	P22413 h ectonucle
4	2434	50.7	906	1 NPP1_MOUSE	P06802 m ectonucle
5	2405	50.1	906	1 NPP1_RAT	Q924c3 r ectonucle
6	2111	43.9	862	1 NPP2_MOUSE	Q9186 m ectonucle
7	2091.5	43.5	863	1 NPP2_HUMAN	Q13822 h ectonucle
8	2043.5	42.5	885	1 NPP2_RAT	P25353 saccharomyc
9	434.5	9.0	742	1 YCR6_YEAST	P33997 saccharomyc
10	412.5	8.6	493	1 YEB6_YEAST	P15396 bos taurus
11	301	6.3	61	1 PPDI_BOVIN	Q91jw2 homo sapien
12	130	2.7	476	1 TNAG_HUMAN	P21128 homo sapien
13	129	2.7	369	1 PP11_HUMAN	Q82d32 versinia pe
14	126	2.6	2004	1 YP73_YERPE	P98157 gallus gall
15	121	2.5	4543	1 LRPI_CHICK	Q91f83 arabisdopsis
16	119	2.5	1378	1 WR52_ARATH	Q09782 schizosacch
17	116.5	2.4	758	1 YAG3_SCHPO	P20613 bombyx mori
18	115.5	2.4	704	1 SSP2_BOMMO	P40798 drosophila
19	114	2.4	1106	1 STC_DROME	Q07954 homo sapien
20	114	2.4	4544	1 LRPI_HUMAN	Q9y2c4 homo sapien
21	113.5	2.4	368	1 NUGL_HUMAN	P10039 gallus gall
22	113.5	2.4	1808	1 TENA_CHICK	Q60793 mus musculu
23	113	2.4	474	1 KLF4_MOUSE	Q10480 schizosacch
24	112.5	2.3	322	1 PNUI_SCHPO	P35950 cricetus
25	112.5	2.3	854	1 LDLR_CFRIG	Q9uul4 schizosacch
26	112.5	2.3	2352	1 MOKC_SCHPO	P08466 saccharomyc
27	112	2.3	879	1 LDLR_RAT	P08466 saccharomyc
28	111	2.3	329	1 NUCI_YEAST	P10675 schistocerc
29	111	2.3	662	1 FAS1_SCHAM	Q9pqb4 ureaplasma
30	111	2.3	1442	1 DP03_UREPA	P03376 chironomus
31	110	2.3	1700	1 BAR3_CHITE	P09849 oryctolagus
32	110	2.3	1926	1 LPH_RABIT	Q81w6 homo sapien
33	109.5	2.3	871	1 SULLI_HUMAN	

#### RESULT 1

NPP3\_HUMAN

ID NPP3\_HUMAN STANDARD; PRT; 875 AA.

AC O14638;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Ectonucleotide pyrophosphatase/phosphodiesterase 3 (E-NPP 3)

DE (Phosphodiesterase I/nucleotide pyrophosphatase 3) (Phosphodiesterase

I beta) (PD-beta) (CD203c antigen) [includes: Alkaline

DE phosphodiesterase I (EC 3.1.4.1); Nucleotide pyrophosphatase

DE (EC 3.6.1.9) (NPPase)].

GN ENPP3 OR PDNP3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;

RX MEDLINE=98008933; PubMed=9344668;

RA Piao J.-H., Goding J.W., Nakamura H., Sano K.;

RT "Molecular cloning and chromosomal localization of PD-beta (PDNP3), a

new member of the human phosphodiesterase I genes.";

[2]

SEQUENCE OF 189-875 FROM N.A.

Hou S., Wohldmann P., Le T.;

Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: Cleaves a variety of phosphodiester and phosphosulfate

bonds including deoxynucleotides, nucleotide sugars, and NAD (By

similarity).

-!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides

successively from the 3'-hydroxy termini of 3'-hydroxy-terminated

oligo-nucleotides.

-!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.

-!- SUBCELLULAR LOCATION: Type II membrane protein.

-!- SIMILARITY: Contains 2 somatomedin-B type domains.

-!- DATABASE: NAME=PROW; NOTE=PROW 1:47-49(2000);

WWW="http://www.ncbi.nlm.nih.gov/prow/guide/575821061.g.htm".

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

-----

EMBL; AF005632; AAC51813.1;

EMBL; AC005587; AAD05192.1;

Genew; HGNC:3358; ENPP3.

MTM; 602182; --

GO; GO:0005887; C:integral to plasma membrane; TAS.

GO; GO:0004551; F:nucleotide diphosphatase activity; TAS.

GO; GO:0004528; F:phosphodiesterase I activity; TAS.

GO; GO:0004528; F:phosphodiesterase I activity; TAS.

GO; GO:0004528; F:phosphodiesterase I activity; TAS.

GO; GO:0004528; F:phosphodiesterase I activity; TAS.





[illegible]

Qy	180	LMPEINLKLTGCIHSKYRWAYPTKPNPHNYHTVLTGLYPESHGIDNNMYDVNLKNFSL	233
Db	181	LLPENLKLTGCLHSKYRWAYPTKPNPHNYHTVLTGLYPESHGIDNNMYDVNLKNFSL	240
Qy	240	SSKEONFAWHGQPMMLTAMYQGLKAATYFWPGSEVAINGSPSTIYPYNGSVPEERI	299
Db	241	SSVEKSNFAWWSGPILWTAMYQGLKAASYVWPGSDVAVNGSPFNIRYNSVPEYERI	300
Qy	300	STLLKWLDPKAEPRPYTWYFEPDSSGHAGGPGVSARVTKALQVVDHAFGLMEGLKOR	359
Db	301	ATLLQWLDPKAEPRPYTWYFEPDSSGHAGGPGVSAGVTKALQVVDHAFGLMEGLKOR	360
Qy	360	NLHNCVNIILLADHGMDOITYCNKMEYMTDFPRINFYMYEGPAPRIRAHNIPHDFFSN	419
Db	361	NLHNCVNIILLADHGMDOITCDRVYMYMTDFPEIN-FYMYQGPAPRIRATENIPQDFFTN	419
Qy	420	SEEVRLNLSCKPQHQFKPYLTDPDLKRLHYAKNVRIDKVLHFPVQOVLAVRSKNTWCG	479
Db	420	SEEVRLNLSCKPQHQFKPYLTDPDLKRLHYAKNVRIDKVLHFPVQOVLAVRSKNTWCG	479
Qy	480	GNHGYNNEFFSMGAI FLAHGSPSKEKTEVEPENTEVNLMCDLLRIOPAPNNGSHGSL	539
Db	480	GTHGYNNEFFSMGAI FLAHGSPSKEKTEVEPENTEVNLMCDLLRIOPAPNNGSHGSL	539
Qy	540	NHLKVPYFPESHAEEVSKFSCVGFANPLPTESLDCFCPLHNSLTQLEQVNMQLNLTQEE	599
Db	540	NHLKVPYFPESHAEEVSKFSCVGFANPLPTESLDCFCPLHNSLTQLEQVNMQLNLTQEE	599
Qy	600	ITATYKVLNPGRRPRVLQKNDHCLLYHREVSVGFGKAMPMWSSVTPVPLGDTSLPLP	659
Db	600	VSATEKTNLPGRRPRVLQKNDHCLLYHREVSVGFGKAMPMWSSVTPVPLGDTSLPLP	659
Qy	660	TVPDCLRADVRVPSESQKCSFYLLADKNITHGFLYPPASNRTSDSQYDALITSLNLYPME	719
Db	660	TVPDCLRADVRVPSESQKCSFYLLADKNITHGFLYPPASNRTSDSQYDALITSLNLYPME	719
Qy	720	EFKRWMDYFHSVLLIKHATERNGVNVSGPIFDYNDYDGHDPADDEITKHLANDVPIPH	779
Db	720	EFKRWMDYFHSVLLIKHATERNGVNVSGPIFDYNDYDGHDPADDEITKHLANDVPIPH	779
Qy	780	YFVWLTSCKNKSHTPCNPGMLDVLPTIIPHRPNVSVESCEGKPEALWVEERTAHIAV	839
Db	780	YFVWLTSCKNKSHTPCNPGMLDVLPTIIPHRPNVSVESCEGKPEALWVEERTAHIAV	839
Qy	840	RDVELLTGLDFYQDKVQPVSEILOLKYLPFTETI 875	
Db	840	RDVELLTGLDFYQDKVQPVSEILOLKYLPFTETI 875	

RESULT 3

NPPI\_HUMAN STANDARD; PR7; 925 AA.

ID NPPI\_HUMAN ID P22413; Q9NP61; Q9U61; Q9Y6K3;

AC P22413; Q9NP61; Q9U61; Q9Y6K3;

DT 01-AUG-1991 (Rel. 19, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)

DE Phosphodiesterase I/nucleotide pyrophosphatase 1 (Plasma-cell

DE membrane glycoprotein PC-1) [Includes: Alkaline phosphodiesterase I

DE (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.5) (NPPase)].

GN ENPP1 OR PNPP1 OR PC1 OR NPPs.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP MEDLINE=91009202; PubMed=2211644;

RX Buckley M.F., Loveland K.A., McKinstry W.J., Garson O.M., Goding J.W.;

RA "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human

FT molecule, amino acid sequence, and chromosomal location.";

RL J. Biol. Chem. 265:17506-17511(1990).

RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=92246539; PubMed=1315502;  
RA Funakoshi I., Kato H., Horie K., Yano T., Hori Y., Kobayashi H.,  
RA Inoue T., Suzuki H., Fukui S., Tsukahara M., Kajii T.,  
RA Yamaehina I.,  
RT "Molecular cloning of cDNAs for human fibroblast nucleotide  
RT pyrophosphatase.";  
RA Arch. Biochem. Biophys. 295:180-187(1992).  
RN [3].  
RP SEQUENCE FROM N.A.  
RA Bozzali M., Pizzuti A., Trischitta E.,  
RT "Genomic structure of the human PC1 gene.";  
RN Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [4].  
RP SEQUENCE OF 81-925 FROM N.A.  
RA Peck A.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [5].  
RP SEQUENCE OF 145-185 FROM N.A., AND VARIANT GLN-173.  
RX MEDLINE=99408501; PubMed=10480624;  
RA Pizzuti A., Frittitta L., Argiolas A., Baratta R., Goldfine I.D.,  
RA Bozzali M., Ercolino T., Scarlato G., Iacoviello L., Vigneri R.,  
RA Tassi V., Trischitta E.,  
RT "A polymorphism (X121Q) of the human glycoprotein PC-1 gene coding  
RT region is strongly associated with insulin resistance.";  
RL Diabetes 48:1881-1884(1999).  
RN [6].  
RP CHARACTERIZATION.  
RX MEDLINE=95094801; PubMed=8001561;  
RA Belli S.I., Goding J.W.,  
RT "Biochemical characterization of human PC-1, an enzyme possessing  
RT alkaline phosphodiesterase I and nucleotide pyrophosphatase  
RT activities.";  
RL Eur. J. Biochem. 226:433-443(1994).  
RN [7].  
RP ACTIVE SITE  
RX MEDLINE=95255270; PubMed=7737162;  
RA Belli S.I., Mercuri F.A., Sali A., Goding J.W.,  
RT "Autophosphorylation of PC-1 (alkaline phosphodiesterase I/nucleotide  
RT pyrophosphatase) and analysis of the active site.";  
RL Eur. J. Biochem. 228:669-676(1995).  
RN [8].  
RP VARIANTS OPLL PRO-91 AND PHE-287, AND VARIANTS GLN-173; HIS-268 AND  
RP PRO-779.  
RX MEDLINE=99381541; PubMed=10453738;  
RA Nakamura I., Ikegawa S., Okawa A., Okuda S., Koshizuka Y.,  
RA Kawaguchi H., Nakamura K., Koyama T., Goto S., Toguchida J.,  
RA Matsushita M., Ochi T., Takaoka K., Nakamura Y.,  
RT "Association of the human NPPS gene with ossification of the posterior  
RT longitudinal ligament of the spine (OPLL).";  
RL Hum. Genet. 104:492-497(1999).  
RN [9].  
RP FUNCTION: Has a broad specificity and cleaves a variety of  
CC substrates, including phosphodiester bonds of nucleotides and  
CC nucleotide sugars and pyrophosphate bonds of nucleotides and  
CC nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such  
CC as ATP, GTP, CTP, TTP and UTP to their corresponding  
CC monophosphates with release of pyrophosphate. Can hydrolyze  
CC diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a  
CC role in the regulation of pyrophosphate production, the regulation  
CC of the availability of nucleotide sugars in the endoplasmic  
CC reticulum and Golgi, and the regulation of purinergic signaling.  
CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides  
CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated  
CC oligo-nucleotides.  
CC -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.  
CC -!- SUBUNIT: Homodimer; disulfide-linked.  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER  
CC OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE  
CC OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.  
CC -!- PTM: Autophosphorylated as part of the catalytic cycle of  
CC phosphodiesterase/pyrophosphatase activity.  
CC -!- PTM: N-glycosylated.

CC -!- DISEASE: Defects in ENPPI are a cause of increased susceptibility  
CC for ossification of the posterior longitudinal ligament of the  
CC spine (OPLL), a common form of human myelopathy with a prevalence  
CC of as much as 4% in a variety of ethnic groups.  
CC -!- SIMILARITY: Contains 2 somatomedin-B type domains.  
CC -!- DATABASE: NAME=PRO; NOTE=PROW 2:32-44(2001);  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/31658143\_g.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M57736; AAA63237.1; ALT INIT.  
DR EMBL; D12485; BAA02054.1; ALT INIT.  
DR EMBL; AF110304; AAF36094.1; -  
DR EMBL; AF110280; AAF36094.1; JOINED.  
DR EMBL; AF110281; AAF36094.1; JOINED.  
DR EMBL; AF110283; AAF36094.1; JOINED.  
DR EMBL; AF110284; AAF36094.1; JOINED.  
DR EMBL; AF110285; AAF36094.1; JOINED.  
DR EMBL; AF110286; AAF36094.1; JOINED.  
DR EMBL; AF110287; AAF36094.1; JOINED.  
DR EMBL; AF110288; AAF36094.1; JOINED.  
DR EMBL; AF110289; AAF36094.1; JOINED.  
DR EMBL; AF110290; AAF36094.1; JOINED.  
DR EMBL; AF110291; AAF36094.1; JOINED.  
DR EMBL; AF110292; AAF36094.1; JOINED.  
DR EMBL; AF110293; AAF36094.1; JOINED.  
DR EMBL; AF110294; AAF36094.1; JOINED.  
DR EMBL; AF110295; AAF36094.1; JOINED.  
DR EMBL; AF110296; AAF36094.1; JOINED.  
DR EMBL; AF110297; AAF36094.1; JOINED.  
DR EMBL; AF110298; AAF36094.1; JOINED.  
DR EMBL; AF110299; AAF36094.1; JOINED.  
DR EMBL; AF110300; AAF36094.1; JOINED.  
DR EMBL; AF110301; AAF36094.1; JOINED.  
DR EMBL; AF110302; AAF36094.1; JOINED.  
DR EMBL; AF110303; AAF36094.1; JOINED.  
DR EMBL; AF110304; AAF36094.1; JOINED.  
DR EMBL; AF110305; AAF36094.1; JOINED.  
DR EMBL; AF110306; AAF36094.1; JOINED.  
DR EMBL; AF110307; AAF36094.1; JOINED.  
DR EMBL; AF110308; AAF36094.1; JOINED.  
DR EMBL; AF110309; AAF36094.1; JOINED.  
DR EMBL; AF110310; AAF36094.1; JOINED.  
DR EMBL; AF110311; AAF36094.1; JOINED.  
DR EMBL; AF110312; AAF36094.1; JOINED.  
DR EMBL; AF110313; AAF36094.1; JOINED.  
DR EMBL; AF110314; AAF36094.1; JOINED.  
DR EMBL; AF110315; AAF36094.1; JOINED.  
DR EMBL; AF110316; AAF36094.1; JOINED.  
DR EMBL; AF110317; AAF36094.1; JOINED.  
DR EMBL; AF110318; AAF36094.1; JOINED.  
DR EMBL; AF110319; AAF36094.1; JOINED.  
DR EMBL; AF110320; AAF36094.1; JOINED.  
DR EMBL; AF110321; AAF36094.1; JOINED.  
DR EMBL; AF110322; AAF36094.1; JOINED.  
DR EMBL; AF110323; AAF36094.1; JOINED.  
DR EMBL; AF110324; AAF36094.1; JOINED.  
DR EMBL; AF110325; AAF36094.1; JOINED.  
DR EMBL; AF110326; AAF36094.1; JOINED.  
DR EMBL; AF110327; AAF36094.1; JOINED.  
DR EMBL; AF110328; AAF36094.1; JOINED.  
DR EMBL; AF110329; AAF36094.1; JOINED.  
DR EMBL; AF110330; AAF36094.1; JOINED.  
DR EMBL; AF110331; AAF36094.1; JOINED.  
DR EMBL; AF110332; AAF36094.1; JOINED.  
DR EMBL; AF110333; AAF36094.1; JOINED.  
DR EMBL; AF110334; AAF36094.1; JOINED.  
DR EMBL; AF110335; AAF36094.1; JOINED.  
DR EMBL; AF110336; AAF36094.1; JOINED.  
DR EMBL; AF110337; AAF36094.1; JOINED.  
DR EMBL; AF110338; AAF36094.1; JOINED.  
DR EMBL; AF110339; AAF36094.1; JOINED.  
DR EMBL; AF110340; AAF36094.1; JOINED.  
DR EMBL; AF110341; AAF36094.1; JOINED.  
DR EMBL; AF110342; AAF36094.1; JOINED.  
DR EMBL; AF110343; AAF36094.1; JOINED.  
DR EMBL; AF110344; AAF36094.1; JOINED.  
DR EMBL; AF110345; AAF36094.1; JOINED.  
DR EMBL; AF110346; AAF36094.1; JOINED.  
DR EMBL; AF110347; AAF36094.1; JOINED.  
DR EMBL; AF110348; AAF36094.1; JOINED.  
DR EMBL; AF110349; AAF36094.1; JOINED.  
DR EMBL; AF110350; AAF36094.1; JOINED.  
DR EMBL; AF110351; AAF36094.1; JOINED.  
DR EMBL; AF110352; AAF36094.1; JOINED.  
DR EMBL; AF110353; AAF36094.1; JOINED.  
DR EMBL; AF110354; AAF36094.1; JOINED.  
DR EMBL; AF110355; AAF36094.1; JOINED.  
DR EMBL; AF110356; AAF36094.1; JOINED.  
DR EMBL; AF110357; AAF36094.1; JOINED.  
DR EMBL; AF110358; AAF36094.1; JOINED.  
DR EMBL; AF110359; AAF36094.1; JOINED.  
DR EMBL; AF110360; AAF36094.1; JOINED.  
DR EMBL; AF110361; AAF36094.1; JOINED.  
DR EMBL; AF110362; AAF36094.1; JOINED.  
DR EMBL; AF110363; AAF36094.1; JOINED.  
DR EMBL; AF110364; AAF36094.1; JOINED.  
DR EMBL; AF110365; AAF36094.1; JOINED.  
DR EMBL; AF110366; AAF36094.1; JOINED.  
DR EMBL; AF110367; AAF36094.1; JOINED.  
DR EMBL; AF110368; AAF36094.1; JOINED.  
DR EMBL; AF110369; AAF36094.1; JOINED.  
DR EMBL; AF110370; AAF36094.1; JOINED.  
DR EMBL; AF110371; AAF36094.1; JOINED.  
DR EMBL; AF110372; AAF36094.1; JOINED.  
DR EMBL; AF110373; AAF36094.1; JOINED.  
DR EMBL; AF110374; AAF36094.1; JOINED.  
DR EMBL; AF110375; AAF36094.1; JOINED.  
DR EMBL; AF110376; AAF36094.1; JOINED.  
DR EMBL; AF110377; AAF36094.1; JOINED.  
DR EMBL; AF110378; AAF36094.1; JOINED.  
DR EMBL; AF110379; AAF36094.1; JOINED.  
DR EMBL; AF110380; AAF36094.1; JOINED.  
DR EMBL; AF110381; AAF36094.1; JOINED.  
DR EMBL; AF110382; AAF36094.1; JOINED.  
DR EMBL; AF110383; AAF36094.1; JOINED.  
DR EMBL; AF110384; AAF36094.1; JOINED.  
DR EMBL; AF110385; AAF36094.1; JOINED.  
DR EMBL; AF110386; AAF36094.1; JOINED.  
DR EMBL; AF110387; AAF36094.1; JOINED.  
DR EMBL; AF110388; AAF36094.1; JOINED.  
DR EMBL; AF110389; AAF36094.1; JOINED.  
DR EMBL; AF110390; AAF36094.1; JOINED.  
DR EMBL; AF110391; AAF36094.1; JOINED.  
DR EMBL; AF110392; AAF36094.1; JOINED.  
DR EMBL; AF110393; AAF36094.1; JOINED.  
DR EMBL; AF110394; AAF36094.1; JOINED.  
DR EMBL; AF110395; AAF36094.1; JOINED.  
DR EMBL; AF110396; AAF36094.1; JOINED.  
DR EMBL; AF110397; AAF36094.1; JOINED.  
DR EMBL; AF110398; AAF36094.1; JOINED.  
DR EMBL; AF110399; AAF36094.1; JOINED.  
DR EMBL; AF110400; AAF36094.1; JOINED.  
DR EMBL; AF110401; AAF36094.1; JOINED.  
DR EMBL; AF110402; AAF36094.1; JOINED.  
DR EMBL; AF110403; AAF36094.1; JOINED.  
DR EMBL; AF110404; AAF36094.1; JOINED.  
DR EMBL; AF110405; AAF36094.1; JOINED.  
DR EMBL; AF110406; AAF36094.1; JOINED.  
DR EMBL; AF110407; AAF36094.1; JOINED.  
DR EMBL; AF110408; AAF36094.1; JOINED.  
DR EMBL; AF110409; AAF36094.1; JOINED.  
DR EMBL; AF110410; AAF36094.1; JOINED.  
DR EMBL; AF110411; AAF36094.1; JOINED.  
DR EMBL; AF110412; AAF36094.1; JOINED.  
DR EMBL; AF110413; AAF36094.1; JOINED.  
DR EMBL; AF110414; AAF36094.1; JOINED.  
DR EMBL; AF110415; AAF36094.1; JOINED.  
DR EMBL; AF110416; AAF36094.1; JOINED.  
DR EMBL; AF110417; AAF36094.1; JOINED.  
DR EMBL; AF110418; AAF36094.1; JOINED.  
DR EMBL; AF110419; AAF36094.1; JOINED.  
DR EMBL; AF110420; AAF36094.1; JOINED.  
DR EMBL; AF110421; AAF36094.1; JOINED.  
DR EMBL; AF110422; AAF36094.1; JOINED.  
DR EMBL; AF110423; AAF36094.1; JOINED.  
DR EMBL; AF110424; AAF36094.1; JOINED.  
DR EMBL; AF110425; AAF36094.1; JOINED.  
DR EMBL; AF110426; AAF36094.1; JOINED.  
DR EMBL; AF110427; AAF36094.1; JOINED.  
DR EMBL; AF110428; AAF36094.1; JOINED.  
DR EMBL; AF110429; AAF36094.1; JOINED.  
DR EMBL; AF110430; AAF36094.1; JOINED.  
DR EMBL; AF110431; AAF36094.1; JOINED.  
DR EMBL; AF110432; AAF36094.1; JOINED.  
DR EMBL; AF110433; AAF36094.1; JOINED.  
DR EMBL; AF110434; AAF36094.1; JOINED.  
DR EMBL; AF110435; AAF36094.1; JOINED.  
DR EMBL; AF110436; AAF36094.1; JOINED.  
DR EMBL; AF110437; AAF36094.1; JOINED.  
DR EMBL; AF110438; AAF36094.1; JOINED.  
DR EMBL; AF110439; AAF36094.1; JOINED.  
DR EMBL; AF110440; AAF36094.1; JOINED.  
DR EMBL; AF110441; AAF36094.1; JOINED.  
DR EMBL; AF110442; AAF36094.1; JOINED.  
DR EMBL; AF110443; AAF36094.1; JOINED.  
DR EMBL; AF110444; AAF36094.1; JOINED.  
DR EMBL; AF110445; AAF36094.1; JOINED.  
DR EMBL; AF110446; AAF36094.1; JOINED.  
DR EMBL; AF110447; AAF36094.1; JOINED.  
DR EMBL; AF110448; AAF36094.1; JOINED.  
DR EMBL; AF110449; AAF36094.1; JOINED.  
DR EMBL; AF110450; AAF36094.1; JOINED.  
DR EMBL; AF110451; AAF36094.1; JOINED.  
DR EMBL; AF110452; AAF36094.1; JOINED.  
DR EMBL; AF110453; AAF36094.1; JOINED.  
DR EMBL; AF110454; AAF36094.1; JOINED.  
DR EMBL; AF110455; AAF36094.1; JOINED.  
DR EMBL; AF110456; AAF36094.1; JOINED.  
DR EMBL; AF110457; AAF36094.1; JOINED.  
DR EMBL; AF110458; AAF36094.1; JOINED.  
DR EMBL; AF110459; AAF36094.1; JOINED.  
DR EMBL; AF110460; AAF36094.1; JOINED.  
DR EMBL; AF110461; AAF36094.1; JOINED.  
DR EMBL; AF110462; AAF36094.1; JOINED.  
DR EMBL; AF110463; AAF36094.1; JOINED.  
DR EMBL; AF110464; AAF36094.1; JOINED.  
DR EMBL; AF110465; AAF36094.1; JOINED.  
DR EMBL; AF110466; AAF36094.1; JOINED.  
DR EMBL; AF110467; AAF36094.1; JOINED.  
DR EMBL; AF110468; AAF36094.1; JOINED.  
DR EMBL; AF110469; AAF36094.1; JOINED.  
DR EMBL; AF110470; AAF36094.1; JOINED.  
DR EMBL; AF110471; AAF36094.1; JOINED.  
DR EMBL; AF110472; AAF36094.1; JOINED.  
DR EMBL; AF110473; AAF36094.1; JOINED.  
DR EMBL; AF110474; AAF36094.1; JOINED.  
DR EMBL; AF110475; AAF36094.1; JOINED.  
DR EMBL; AF110476; AAF36094.1; JOINED.  
DR EMBL; AF110477; AAF36094.1; JOINED.  
DR EMBL; AF110478; AAF36094.1; JOINED.  
DR EMBL; AF110479; AAF36094.1; JOINED.  
DR EMBL; AF110480; AAF36094.1; JOINED.  
DR EMBL; AF110481; AAF36094.1; JOINED.  
DR EMBL; AF110482; AAF36094.1; JOINED.  
DR EMBL; AF110483; AAF36094.1; JOINED.  
DR EMBL; AF110484; AAF36094.1; JOINED.  
DR EMBL; AF110485; AAF36094.1; JOINED.  
DR EMBL; AF110486; AAF36094.1; JOINED.  
DR EMBL; AF110487; AAF36094.1; JOINED.  
DR EMBL; AF110488; AAF36094.1; JOINED.  
DR EMBL; AF110489; AAF36094.1; JOINED.  
DR EMBL; AF110490; AAF36094.1; JOINED.  
DR EMBL; AF110491; AAF36094.1; JOINED.  
DR EMBL; AF110492; AAF36094.1; JOINED.  
DR EMBL; AF110493; AAF36094.1; JOINED.  
DR EMBL; AF110494; AAF36094.1; JOINED.  
DR EMBL; AF110495; AAF36094.1; JOINED.  
DR EMBL; AF110496; AAF36094.1; JOINED.  
DR EMBL; AF110497; AAF36094.1; JOINED.  
DR EMBL; AF110498; AAF36094.1; JOINED.  
DR EMBL; AF110499; AAF36094.1; JOINED.  
DR EMBL; AF110500; AAF36094.1; JOINED.  
DR EMBL; AF110501; AAF36094.1; JOINED.  
DR EMBL; AF110502; AAF36094.1; JOINED.  
DR EMBL; AF110503; AAF36094.1; JOINED.  
DR EMBL; AF110504; AAF36094.1; JOINED.  
DR EMBL; AF110505; AAF36094.1; JOINED.  
DR EMBL; AF110506; AAF36094.1; JOINED.  
DR EMBL; AF110507; AAF36094.1; JOINED.  
DR EMBL; AF110508; AAF36094.1; JOINED.  
DR EMBL; AF110509; AAF36094.1; JOINED.  
DR EMBL; AF110510; AAF36094.1; JOINED.  
DR EMBL; AF110511; AAF36094.1; JOINED.  
DR EMBL; AF110512; AAF36094.1; JOINED.  
DR EMBL; AF110513; AAF36094.1; JOINED.  
DR EMBL; AF110514; AAF36094.1; JOINED.  
DR EMBL; AF110515; AAF36094.1; JOINED.  
DR EMBL; AF110516; AAF36094.1; JOINED.  
DR EMBL; AF110517; AAF36094.1; JOINED.  
DR EMBL; AF110518; AAF36094.1; JOINED.  
DR EMBL; AF110519; AAF36094.1; JOINED.  
DR EMBL; AF110520; AAF36094.1; JOINED.  
DR EMBL; AF110521; AAF36094.1; JOINED.  
DR EMBL; AF110522; AAF36094.1; JOINED.  
DR EMBL; AF110523; AAF36094.1; JOINED.  
DR EMBL; AF110524; AAF36094.1; JOINED.  
DR EMBL; AF110525; AAF36094.1; JOINED.  
DR EMBL; AF110526; AAF36094.1; JOINED.  
DR EMBL; AF110527; AAF36094.1; JOINED.  
DR EMBL; AF110528; AAF36094.1; JOINED.  
DR EMBL; AF110529; AAF36094.1; JOINED.  
DR EMBL; AF110530; AAF36094.1; JOINED.  
DR EMBL; AF110531; AAF36094.1; JOINED.  
DR EMBL; AF110532; AAF36094.1; JOINED.  
DR EMBL; AF110533; AAF36094.1; JOINED.  
DR EMBL; AF110534; AAF36094.1; JOINED.  
DR EMBL; AF110535; AAF36094.1; JOINED.  
DR EMBL; AF110536; AAF36094.1; JOINED.  
DR EMBL; AF110537; AAF36094.1; JOINED.  
DR EMBL; AF110538; AAF36094.1; JOINED.  
DR EMBL; AF110539; AAF36094.1; JOINED.  
DR EMBL; AF110540; AAF36094.1; JOINED.  
DR EMBL; AF110541; AAF36094.1; JOINED.  
DR EMBL; AF110542; AAF36094.1; JOINED.  
DR EMBL; AF110543; AAF36094.1; JOINED.  
DR EMBL; AF110544; AAF36094.1; JOINED.  
DR EMBL; AF110545; AAF36094.1; JOINED.  
DR EMBL; AF110546; AAF36094.1; JOINED.  
DR EMBL; AF110547; AAF36094.1; JOINED.  
DR EMBL; AF110548; AAF36094.1; JOINED.  
DR EMBL; AF110549; AAF36094.1; JOINED.  
DR EMBL; AF110550; AAF36094.1; JOINED.  
DR EMBL; AF110551; AAF36094.1; JOINED.  
DR EMBL; AF110552; AAF36094.1; JOINED.  
DR EMBL; AF110553; AAF36094.1; JOINED.  
DR EMBL; AF110554; AAF36094.1; JOINED.  
DR EMBL; AF110555; AAF36094.1; JOINED.  
DR EMBL; AF110556; AAF36094.1; JOINED.  
DR EMBL; AF110557; AAF36094.1; JOINED.  
DR EMBL; AF110558; AAF36094.1; JOINED.  
DR EMBL; AF110559; AAF36094.1; JOINED.  
DR EMBL; AF110560; AAF36094.1; JOINED.  
DR EMBL; AF110561; AAF36094.1; JOINED.  
DR EMBL; AF110562; AAF36094.1; JOINED.  
DR EMBL; AF110563; AAF36094.1; JOINED.  
DR EMBL; AF110564; AAF36094.1; JOINED.  
DR EMBL; AF110565; AAF36094.1; JOINED.  
DR EMBL; AF110566; AAF36094.1; JOINED.  
DR EMBL; AF110567; AAF36094.1; JOINED.  
DR EMBL; AF110568; AAF36094.1; JOINED.  
DR EMBL; AF110569; AAF36094.1; JOINED.  
DR EMBL; AF110570; AAF36094.1; JOINED.  
DR EMBL; AF110571; AAF36094.1; JOINED.  
DR EMBL; AF110572; AAF36094.1; JOINED.  
DR EMBL; AF110573; AAF36094.1; JOINED.  
DR EMBL; AF110574; AAF36094.1; JOINED.  
DR EMBL; AF110575; AAF36094.1; JOINED.  
DR EMBL; AF110576; AAF36094.1; JOINED.  
DR EMBL; AF110577; AAF36094.1; JOINED.  
DR EMBL; AF110578; AAF36094.1; JOINED.  
DR EMBL; AF110579; AAF36094.1; JOINED.  
DR EMBL; AF110580; AAF36094.1; JOINED.  
DR EMBL; AF110581; AAF36094.1; JOINED.  
DR EMBL; AF110582; AAF36094.1; JOINED.  
DR EMBL; AF110583; AAF36094.1; JOINED.  
DR EMBL; AF110584; AAF36094.1; JOINED.  
DR EMBL; AF110585; AAF36094.1; JOINED.  
DR EMBL; AF110586; AAF36094.1; JOINED.  
DR EMBL; AF110587; AAF36094.1; JOINED.  
DR EMBL; AF110588; AAF36094.1; JOINED.  
DR EMBL; AF110589; AAF36094.1; JOINED.  
DR EMBL; AF110590; AAF36094.1; JOINED.  
DR EMBL; AF110591; AAF36094.1; JOINED.  
DR EMBL; AF110592; AAF36094.1; JOINED.  
DR EMBL; AF110593; AAF36094.1; JOINED.  
DR EMBL; AF110594; AAF36094.1; JOINED.  
DR EMBL; AF110595; AAF36094.1; JOINED.  
DR EMBL; AF110596; AAF36094.1; JOINED.  
DR EMBL; AF110597; AAF36094.1; JOINED.  
DR EMBL; AF110598; AAF36094.1; JOINED.  
DR EMBL; AF110599; AAF36094.1; JOINED.  
DR EMBL; AF110600; AAF36094.1; JOINED.  
DR EMBL; AF110601; AAF36094.1; JOINED.  
DR EMBL; AF110602; AAF36094.1; JOINED.  
DR EMBL; AF110603; AAF36094.1; JOINED.  
DR EMBL; AF110604; AAF36094.1; JOINED.  
DR EMBL; AF110605; AAF36094.1; JOINED.  
DR EMBL; AF110606; AAF36094.1; JOINED.  
DR EMBL; AF110607; AAF36094.1; JOINED.  
DR EMBL; AF110608; AAF36094.1; JOINED.  
DR EMBL; AF110609; AAF36094.1; JOINED.  
DR EMBL; AF110610; AAF36094.1; JOINED.  
DR EMBL; AF110611; AAF36094.1; JOINED.  
DR EMBL; AF110612; AAF36094.1; JOINED.  
DR EMBL; AF110613; AAF36094.1; JOINED.  
DR EMBL; AF110614; AAF36094.1; JOINED.  
DR EMBL; AF110615; AAF36094.1; JOINED.  
DR EMBL; AF110616; AAF36094.1; JOINED.  
DR EMBL; AF110617; AAF36094.1; JOINED.  
DR EMBL; AF110618; AAF36094.1; JOINED.  
DR EMBL; AF110619; AAF36094.1; JOINED.  
DR EMBL; AF110620; AAF36094.1; JOINED.  
DR EMBL; AF110621; AAF36094.1; JOINED.  
DR EMBL; AF110622; AAF36094.1; JOINED.  
DR EMBL; AF110623; AAF36094.1; JOINED.  
DR EMBL; AF110624; AAF36094.1; JOINED.  
DR EMBL; AF110625; AAF36094.1; JOINED.  
DR EMBL; AF110626; AAF36094.1; JOINED.  
DR EMBL; AF110627; AAF36094.1; JOINED.  
DR EMBL; AF110628; AAF36094.1; JOINED.  
DR EMBL; AF110629; AAF36094.1; JOINED.  
DR EMBL; AF110630; AAF36094.1; JOINED.  
DR EMBL; AF110631; AAF36094.1; JOINED.  
DR EMBL; AF110632; AAF36094.1; JOINED.  
DR EMBL; AF110633; AAF36094.1; JOINED.  
DR EMBL; AF110634; AAF36094.1; JOINED.  
DR EMBL; AF110635; AAF36094.1; JOINED.  
DR EMBL; AF110636; AAF36094.1; JOINED.  
DR EMBL; AF110637; AAF36094.1; JOINED.  
DR EMBL; AF110638; AAF36094.1; JOINED.  
DR EMBL; AF110639; AAF36094.1; JOINED.  
DR EMBL; AF110640; AAF36094.1; JOINED.  
DR EMBL; AF110641; AAF36094.1; JOINED.  
DR EMBL; AF110642; AAF36094.1; JOINED.  
DR EMBL; AF110643; AAF36094.1; JOINED.  
DR EMBL; AF110644; AAF36094.1; JOINED.  
DR EMBL; AF110645; AAF36094.1; JOINED.  
DR EMBL; AF110646; AAF36094.1; JOINED.  
DR EMBL; AF110647; AAF36094.1; JOINED.  
DR EMBL; AF110648; AAF36094.1; JOINED.  
DR EMBL; AF110649; AAF36094.1; JOINED.  
DR EMBL; AF110650; AAF36094.1; JOINED.  
DR EMBL; AF110651; AAF36094.1; JOINED.  
DR EMBL; AF110652; AAF36094.1; JOINED.  
DR

DR MIM; 602475; -. C: integral to membrane; TAS.  
DR GO; GO:0016021; F: nucleotide diphosphate activity; TAS.  
DR GO; GO:0004551; F: energy pathways; TAS.  
DR GO; GO:0006091; P: phosphate metabolism; TAS.  
DR GO; GO:0007936; P: response to nutrients; TAS.  
DR GO; GO:0007584; P: response to nutrients; TAS.  
DR InterPro; IPR001604; Endonuclease.  
DR InterPro; IPR002591; Phosphodiesterase.  
DR InterPro; IPR001212; Somatomedin\_B.  
DR Pfam; PF01223; Endonuclease; 1.  
DR Pfam; PF01663; Phosphodiesterase; 1.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR SMART; SM00477; NUC; 1.  
DR SMART; SM00201; SO; 2.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
KW Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;  
KW Multifunctional enzyme; Polymorphism; Disease mutation;  
KW Phosphorylation.  
FT DOMAIN 1 76 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 77 97 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
FT FT 98 925 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 104 144 SOMATOMEDIN-B LIKE 1.  
FT DOMAIN 145 188 SOMATOMEDIN-B LIKE 2.  
FT DOMAIN 191 591 PHOSPHODIESTERASE.  
FT DOMAIN 654 925 NUCLEASE.  
FT ACT\_SITE 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).  
Query Match 52.0%; Score 2498.5; DB 1; Length 925;  
Best Local Similarity 53.1%; Pred. No. 4e-171;  
Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;  
QY 10 EQPVKK---NTLK---KYKACIVLLALLVIMSLGLGLGLGLRLKXGGRCKRCKPDAS 62  
DB 57 EPLEKAARARTAKDNTYKVLVLVSVCLTILGCIPLGLKPKSCAKVSKGRCFCFT 116  
QY 63 FRLGNCRDVAKRGDCWFDTCVETRIWNCNRCGCTEALSCSDCDLQK 122  
DB 117 F---GNCRCDAAVEIGNCCLDYQETCIEPIHWTCNRCGCKRLTSLCACSDCKDK 173  
QY 123 KCCADYKVCQGETSWLENCOTAOQCPCGFLDPPVILFPMGFAEYLYTWDTLMP 182  
DB 174 GDCCLNYSVCQGEKSWVEPCESINPCQAGFTPTLLSLDGFRAEYLYHTWGLLP 233  
QY 183 NINKLTCGHSKMYEAMTYTKTFPNHYTIVTGLYPESHGIIIDNMVNLNKNFSLSK 242  
DB 234 VISLKKCGTYTKQMPVYTKTFPNHYSGIVTGLYPESHGIIIDNMVNLNKNFSLSK 293  
QY 243 EQNNPAWHGQPMWTAMYGKLAATYFPGSEVAINGSPSIYMPYNGSVFEEIRISTL 302  
DB 294 EKFNPEWYGEPIWVYAKYQGLKSGTFPFGSDVEINGFPDIYKMYNGSVFEEIRILAV 353  
QY 303 LKWLDPKAEPRFYTMYFEEDSSGHAGGVSARVVKALQVVDHAFGLMGLKORNLH 362  
DB 354 LQWLQPKDERPHFYTLLEEDPSGSHSYGPVSSEVIKALQVVDHAFGLMGLKORNLH 413  
QY 363 NCVNILLADHGMQDYCKNMEYTYDFPRNFYVGEPAEIRAHNIPHPDFFNSSEE 422  
DB 414 RCLNLILISDHGMEQSCCKYIYLNKYLGDVKNKIYVGAALRPSDVPDKYFNBYBG 473  
QY 423 IVNLSCKRQDQHPKPYLTDLPRKRLHYAKNVRIDKVHLFVDQW-LAVRSKNTNCGGG 481  
DB 474 IARNLSREENQHPKPYLKHFLPKRLHFAKSDRIEPLTFYLDLPQWLALNPSEKCYCGG 533  
QY 482 NHGVNNEFRSMEALFLAHGSPFEKTEVEPFENIEVNLNMLCDLLRLOPAPNNGTHGLNH 541  
DB 534 FHGSDNVFSNNQALFVYGPGFKHGIEADTFENIEVNLNMLCDLLNLTAPNNGTHGLNH 593

QY 542 LLKVPFEPESHAEVSKFSCVGFANPLPTESLDQFC-PHLQNSTQLEQVQNMNLTOEEI 600  
DB 594 LLKNFYVTKPKKEVHPVLCQCFTRN-PRDNLGCSNPSI---LPEDPQTQFNLTVAEE 649  
QY 601 TATVKVNLFPGRPVLRQKVDHCLLYHRYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
DB 650 KLIKHTLPYGRPVLRQKENTICLLSQHQFMGYSQDILMPLMTSYVDR--NDSFSTED 707  
QY 661 VPDCLRADVRVPSESKQCSFYLADKNITHGLYPPASNRITDSQY-DALITSNLVPMYE 719  
DB 708 FSNCLYQDFRPLSPVHKSFYKNNTKVSYGFLSPQLNKNSSGIYSEALLTINIVPMYQ 767  
QY 720 EFRKMDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVI 776  
DB 768 SFQVWXYFHDITLLRYAEERNGVNVVSGPVEDFDYDGRCDLSENLQRKRRVIRNOEILI 827  
QY 777 PTHYFVLTSCNKSHTPENCPCWGLDVLFPILPHRTNVVSCPEGKPEALWVEERTTAHI 836  
DB 828 FTHFFIVLTSCKDTSTQPLHCEN-LDTLAFILPHRTDNSESCVHGKHSWVEELMLHR 886  
QY 837 ARVDRVELLTGLDYQDKVPVSEILQLKTYLPTF 871  
DB 887 ARITDVEHITGLSFYQORKEFVSDILKTLHPTF 921  
RESULT 4  
ID\_NPPI\_MOUSE STANDARD; PRT; 906 AA.  
AC P06802;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)  
DE (Phosphodiesterase I/nucleotide pyrophosphatase 1) (Plasma-cell  
DE membrane glycoprotein PC-1) (Iy-41) [includes: Alkaline  
DE phosphodiesterase I (EC 3.1.4.1); Nucleotide pyrophosphatase  
DE (EC 3.6.1.9) (NPPase)].  
DE ENPPI OR PDNPI OR PC1 OR NPPS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALEB/C;  
RX MEDLINE=87165906; PubMed=3104326;  
RA van Driel I.R., Goding J.W.;  
RT "Plasma cell membrane glycoprotein PC-1. Primary structure deduced  
RT from cDNA clones.";  
RL J. Biol. Chem. 262:4882-4887 (1987).  
[2]  
RP REVISIONS TO 24; 46-47; 640-641 AND 692, VARIANTS ARG-651 AND SER-680,  
RP AND ALTERNATIVE SPLICING.  
RX MEDLINE=22117315; PubMed=12121276;  
RA Banakh I., Sali A., Dubljevic V., Grobhen B., Slegers H., Goding J.W.;  
RT "Structural basis of alloypics of ecto-nucleotide  
RT pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein  
RT PC-1) in the mouse and rat, and analysis of allele-specific  
RT xenogeneic antibodies.";  
RL Eur. J. Immunogenet. 29:307-313 (2002).  
[3]  
RP FUNCTION, AND SEQUENCE FROM N.A.  
RX MEDLINE=91271356; PubMed=1647027;  
RA Rebbe N.F., Tong B.D., Finley E.M., Hickman S.;  
RT "Identification of nucleotide pyrophosphatase/alkaline  
RT phosphodiesterase I activity associated with the mouse plasma cell  
RT differentiation antigen PC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5192-5196 (1991).  
[4]  
RP SEQUENCE OF 203-219 FROM N.A.  
RX MEDLINE=80994275; PubMed=3001713;  
RA van Driel I.R., Wilks A.F., Pietersz G.A., Goding J.W.;  
RT "Murine plasma cell membrane antigen PC-1: molecular cloning of cDNA

and analysis of expression.";  
[5]  
Proc. Natl. Acad. Sci. U.S.A. 82:8619-8623 (1985).  
[6]  
MEDLINE=98324779; PubMed=9662402;  
Okawa A., Nakamura I., Goto S., Moriya H., Nakamura Y., Ikegawa S.;  
"Mutation in Npps in a mouse model of ossification of the posterior  
longitudinal ligament of the spine.";  
Nat. Genet. 19:271-273 (1998).  
CC -1- FUNCTION: Has a broad specificity and cleaves a variety of  
substrates, including phosphodiester bonds of nucleotides and  
nucleotide sugars and pyrophosphate bonds of nucleotides and  
nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such  
as ATP, GTP, CTP, TTP and UTP to their corresponding  
monophosphates with release of pyrophosphate. Can hydrolyze  
diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a  
role in the regulation of pyrophosphate production, the regulation  
of the availability of nucleotide sugars in the endoplasmic  
reticulum and Golgi, and the regulation of purinergic signaling  
(By similarity).  
CC -1- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides  
successively from the 3'-hydroxy termini of 3'-hydroxy-terminated  
oligo-nucleotides.  
CC -1- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.  
CC -1- SUBUNIT: Homodimer; disulfide-linked.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=2;  
CC isoId=P06802-1; Sequences=Displayed;  
CC Name=1;  
CC isoId=P06802-2; Sequences=VSP 006748;  
CC -1- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON THE SURFACE OF  
ANTIBODY-SECRETING CELLS.  
CC -1- PTM: Autophosphorylated as part of the catalytic cycle of  
phosphodiesterase/pyrophosphatase activity (By similarity).  
CC -1- PTM: The N-terminus is blocked.  
CC -1- DISEASE: Defects in ENPP1 are the cause of the tiptoe walking  
(ttw) phenotype. Ttw mice exhibit ossification of the spinal  
ligaments.  
CC -1- SIMILARITY: Contains 2 somatomedin-B type domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: J02700; AAA39893.2; -;  
CC EMBL: M12552; AAA39892.1; -;  
CC PIR: A27410; A27410.  
CC MGD: MGI:97370; Enppi.  
CC GO: GO:0001501; P:skeletal development; IMP.  
CC InterPro: IPR001604; Endonuclease.  
CC InterPro: IPR002591; Phosphodiesterase.  
CC InterPro: IPR001212; Somatomedin\_B.  
CC Pfam: PF01223; Endonuclease; 1.  
CC Pfam: PF01663; Phosphodiesterase; 1.  
CC PRINTS: PR00022; Somatomedin\_B; 2.  
CC SMART: SM00477; NUC; 1.  
CC SMART: SM00201; SO; 2.  
CC PROSITE: PS00524; SOMATOMEDIN\_B; 2.  
CC Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;

KW Multifunctional enzyme; Polymorphism; Phosphorylation;  
KW Alternative splicing;  
FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 80 906 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 86 126 SOMATOMEDIN-B LIKE 1.  
FT DOMAIN 127 170 SOMATOMEDIN-B LIKE 2.  
FT DOMAIN 173 573 PHOSPHODIESTERASE.  
FT DOMAIN 535 906 NUCLEASE.  
FT ACT\_SITE 238 238 BY SIMILARITY.  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 630 630 Missing (in isoform 1).  
FT VARIANT 651 651 H -> R (IN ENPP1-B).  
FT VARIANT 680 680 R -> S (IN ENPP1-B).  
SQ SEQUENCE 906 AA; 103175 MW; 068045B0ED0F224D CRC64;  
Query Match 50.7%; Score 2434; DB 1; Length 906;  
Best Local Similarity 51.0%; Pred. No. 1.6e-166;  
Matches 447; Conservative 152; Mismatches 250; Indels 28; Gaps 9;  
QY 10 EQPVKK-----NTLKKYKACIVLALLVMSLGLGLGLKLEKQSGCKKCF 59  
DB 39 EELEKAEARPAKDPNT---YKLSLVLSVCLTILGICIFGLPSCAKVSKNGRCF 95  
QY 60 DASFRGLENCRDVACKRDGDCWDFEDTCVSTRIMCMKFCRGSTRLEASLCSDDC 119  
DB 96 ERTF---SNCRDCAACVSLGNCCLDFQETCEVPTHTWTCNFKRCGRKLSRFVCSACDD 152  
QY 120 LOKKCCADYKSVCOGETSWLSENCDTAQQSCPGDFLPPVILFSMDGFRAYLYTWD 179  
DB 153 KTHNCCINYSVCQDKSWVEETCESIDTPCRAEFESPTLLFSLDGFRAEYLHTGG 212  
QY 180 LMPNINKLTCGIHSKYMEAMPTKTFPNHYITVITGLYPESHGIIDNNYDYNLKNFSL 239  
DB 213 LLPVISKLNCGYTKNMEPTKTFPNHYISVITGLYPESHGIIDNNYDYPKMASFSL 272  
QY 240 SSKQNNPAAWHGQPMWLTAMTQGLKAATYFPGSEVAINGSPFIYMPYNGSVFPEERI 299  
DB 273 KSKKEFNPLWYKQPIWTWTANQEVKSGTYFPGSDVEIDGILPDIYKYNGSVFPEERI 332  
QY 300 STILLKWLDPKAEPRFYTMFEEDPSGSHAGGPVSARVIALQVVDHAFGLMGLKQR 359  
DB 333 LAVLEWLQPSHERPHFYTLYLEEDPSGSHGSPVSSEVIALQKQVLDVGLMDGLKDL 392  
QY 360 NLHNCVNILLADHGMQDTYCNKMEYMTDYPFRINFFYMYEGAPRIAHNIPHDFFSN 419  
DB 393 GLDKCLNLILLSDHGMQGSCKKYVYLNKYLGDVANNVAVYGPAAARLPTDVPETYSFN 452  
QY 420 SEIIVENISCRPDQCHPKYLPDLPKELHYAKNVRIDKVLHFDVDOO--LAVRSKSNINC 478  
DB 453 YEALAKNISCREPNQHFPRYLPKPLKLFHAKSDRIEPLTFYLDQQLALNPSERYK 512  
QY 479 GGGNHGYNNEFRSMEAI FLAHGPSFKETVEPPEFENIEVYLMCDLLRIQAPNNGTHGS 538  
DB 513 GSGFHGSONLFNMQALFYGPAFKHGAEDVSFENIEVYLMCDLLGLIPAPNNGSHGS 572  
QY 539 LNHLLKVPYBESHAEVYSKSVCGFANPLPTESIDCFPHLQNSQLEQVQNLNLQEQ 598  
DB 573 LNHLLKKIYINPSHPKESGFLSQ-----PIKSTNSDLGCTCDPWIVPIKDFEQNLNLT 628  
QY 599 EITATVKVNLFPGRPVTLQKNVDHCLLYHREYVSGFGKAMEPMKWSYTVFQLGTSPLP 658  
DB 629 DVDVDIYHMTVPYGRPIILLKQHHVCLLQQQFLTGYSLLDMLPLWASYTF--LRNDQFSR 686  
QY 659 FTVPCILRADVRVPSESKCSFYIADKNVITHFLYPPASVRTSDSQY-DALITSNLVPM 717

Db 687 DDFNCNLYQDLRIPLSPVHKCSYKSKSLSYGFLTPPLNRVNSHIYSEALLTGNIVPM 746  
 QY 718 YEEPRKWDYFHSYLYLKHATERNGVNVSGPIFDYNDYGHDADEITKH---LANTDV 774  
 Db 747 YQSFVLIWHYLDHTLLQRYAHERENGIVNVSGPFDYDGRYDSLEILKQNSRVIRSQEI 806  
 QY 775 PIPHYFVLTCKNKSHTPCNGMDLVDFIIPHRPNVNSCEGKPEALVZERPTA 834  
 Db 807 LIPHFVLTCKNKSHTPCNGMDLVDFIIPHRPNVNSCEGKPEALVZERPTA 834  
 QY 835 HIARVRDVELLTGLDFYQKQVQVPSVSEILQKLYLPTF 871  
 Db 866 HRAVRTDVELLTGLSFVQDQESVSELLKTHLPF 902

## RESULT 5

NPPI RAT  
 ID NPPI RAT STANDARD; PRT; 906 AA.  
 AC Q924C3; Q920C8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)  
 DE (Phosphodiesterase I/nucleotide pyrophosphatase 1) (Plasma-cell  
 DE membrane glycoprotein PC-1) [includes: Alkaline phosphodiesterase I  
 DE (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].  
 GN ENPP1 OR PDNP1 OR PC1 OR NPPs.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS.  
 RC STRAIN=Lewis, Louvain, and Wistar;  
 RX MEDLINE=22117315; PubMed=12121276;  
 RA Banath I., Sali A., Dubljevic V., Grobden B., Slegers H., Goding J.W.;  
 RT "Structural basis of allotypes of ecto-nucleotide  
 RT pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein  
 RT PC-1) in the mouse and rat, and analysis of allele-specific  
 RT xenogeneic antibodies.";  
 RL Eur. J. Immunogenet. 29:307-313(2002).  
 CC -!- FUNCTION: Has a broad specificity and cleaves a variety of  
 CC substrates, including phosphodiester bonds of nucleotides and  
 CC nucleotide sugars and pyrophosphate bonds of nucleotides and  
 CC nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such  
 CC as ATP, GTP, CTP, TTP and UTP to their corresponding  
 CC monophosphates with release of pyrophosphate. Can hydrolyze  
 CC diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a  
 CC role in the regulation of pyrophosphate production, the regulation  
 CC of the availability of nucleotide sugars in the endoplasmic  
 CC reticulum and Golgi, and the regulation of purinergic signaling  
 CC (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides  
 CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated  
 CC oligo-nucleotides.  
 CC -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Events-Alternative splicing; Named isoforms=2;  
 CC Name=2;  
 CC IsoId=Q924C3-1; Sequence=Displayed;  
 CC Name=1;  
 CC IsoId=Q924C3-2; Sequence=VSP\_006749;  
 CC -!- PTM: Autophosphorylated as part of the catalytic cycle of  
 CC phosphodiesterase/pyrophosphatase activity (By similarity).  
 CC -!- SIMILARITY: Contains 2 somatomedin-B type domains.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF340185; AK69653.1; --  
 CC EMBL; AF340186; AK69654.1; --  
 DR EMBL; AF320054; AL26912.1; --  
 DR InterPro; IPR001604; Endonuclease.  
 DR InterPro; IPR002591; Phosphodiesterase.  
 DR InterPro; IPR01212; Somatomedin B.  
 DR Pfam; PF01223; Endonuclease; 1.  
 DR Pfam; PF01663; Phosphodiesterase; 1.  
 DR Pfam; PF01033; Somatomedin B; 2.  
 DR PRINTS; PR00022; SOMATOMEDINB.  
 DR SMART; SMO0477; NUC; 1.  
 DR SMART; SMO0201; SO; 2.  
 DR PROSITE; PS00524; SOMATOMEDIN B; 2.  
 KW Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;  
 KW Multifunctional enzyme; Polymorphism; Phosphorylation;  
 KW Alternative splicing.  
 FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT DOMAIN 80 906 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 86 126 SOMATOMEDIN-B LIKE 1.  
 FT DOMAIN 127 170 SOMATOMEDIN-B LIKE 2.  
 FT DOMAIN 173 573 PHOSPHODIESTERASE.  
 FT DOMAIN 635 906 NUCLEASE.  
 FT ACT SITE 238 238 BY SIMILARITY.  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 630 630 Missing (in isoform 1).  
 FT VARSPLIC 630 630 /ETID-VSP\_006749  
 FT VARIANT 440 442 RPT -> NPF (IN STRAIN WISTAR).  
 FT VARIANT 457 457 A -> T (IN STRAIN LEW).  
 FT VARIANT 555 555 M -> I (IN STRAIN LEW).  
 FT VARIANT 568 568 E -> G (IN STRAIN WISTAR).  
 FT VARIANT 583 583 T -> V (IN STRAIN LEW).  
 FT VARIANT 592 592 F -> V (IN STRAIN LEW).  
 FT VARIANT 624 624 N -> K (IN STRAIN LEW).  
 FT VARIANT 640 640 N -> H (IN STRAIN LEW).  
 FT VARIANT 774 774 V -> I (IN STRAIN LEW).  
 FT VARIANT 806 806 N -> I (IN STRAIN LEW).  
 FT VARIANT 850 850 T -> I (IN STRAIN LEW).  
 FT VARIANT 898 898 H -> Q (IN STRAIN LEW).  
 SQ SEQUENCE 906 AA; 102942 MW; 71256780B279A919 CRC64;

Query Match 50.1%; Score 2405; DB 1; Length 906;  
 Best Local Similarity 50.9%; Fred. No. 1.9e-164;  
 Matches 445; Conservative 146; Mismatches 261; Indels 22; Gaps 9;

QY 10 EQPVKK---NTLK---KYKIACIVLLALLVIMSLGLGLRLEKQGSCKKCFDAS 62  
 Db 39 EEPLAEARAKTAKDNTYKVLVSLVSVLTILGICFGLKPSCAKEVSKGRCFERT 98  
 QY 63 FRGLNCRCDVAKDCCDCCWDFEETCVESPIWMKFRGCTEFLASLCSDDCLQK 122  
 Db 99 F---SNCRDAAACVSLGNCCLDFFQETCTEPTHTWTCNKRFCGRKLSRFVCSADCKAH 155  
 QY 123 KDCADYKSVCGQETSQWLENCDTAQOQSCPGDFLPVILFMSMDGFAEVLTYWDTLMP 182  
 Db 156 NDCCIYSSVCQEKSKSWEACETIDAPQCPAEFSPPTLLFSLDGFRAEVLHTWGLLP 215  
 QY 183 NINKLTCGHSKYRAMPYPTTFPHNYTIVTGLYPESHGIIIDNNYVNLNKPFLSSK 242  
 Db 216 VTSKLNCKCTYTKNRPVPTTKTFPHNYSIVTGLYPESHGIIIDNNYVNLNKPFLSSK 275  
 QY 243 EQNNPAWHGQPMWLLTAMYGKLAATYFWPGSEVAINGSFFPSIYMPYNGSVPFERISTL 302  
 Db 276 EKENPLWYKQPIWVTANHQEVRSVGTYPWPGSDVEIDGLPDIYKVNGSVPFERILAV 335



QY 303 LKWLDFKAEPRFRFYTYMEFPDSSGHAGQVPSARVVKALQVVDHAFGLMGLKQRLNH 362  
 D 336 LEWLQLPYSRPHFYTYLYLEEDPSSGHSGHPVSSEVVKALQKQVHDVHGLMGLKQDLGLD 395  
 QY 363 NCYNIILLADHGMQDYCNKQKMTYDYPFRINFFNYEGAPRIRAHNIPHPDFFSNSE 422  
 D 396 KCLNLLISDHGMEQSCSKYVYVNLKYLGDVNVKVYVGPAAARLRTPEVTFYVFNSEA 455  
 QY 423 IVENLSCKRQDHFVKVLPDLPKRLHYAKNVRIDKVLHFDVQOM-LAVRSKNTNCGGG 481  
 D 456 LAKNLSRETNQHFYLYLHFKPLHFAKNDRIEPLTFYLPQQLALNPSERKYCGSG 515  
 QY 482 NHGYNFRSMELFLAHGSPFKETVEPPFENIEVYNLMCDLLRIQAPNNGTHGSLNH 541  
 D 516 FHGSDNLFNMQLFTGYGPAFKGAEDVDFENIEVYNLMCDLLGLIPAPNNEHSGSLNH 575  
 QY 542 LKVPFVPEPASHAEVSKSVCGFANPLPTESLDCPCPLHNSQLEQVNMNLNLTQBEIT 601  
 D 576 LKPKPIYTSHPKEESFLQC-----PIKSVSDLGCTCDPSIVPIMDFEKFQFNLTDAVE 631  
 QY 602 ATKVNLPRPRVRLQKNDHCLLYHREYVSGFKAMRPMWSSYTVPLQGDTSPLPPTV 661  
 D 632 DVYSMTVPNGRPNLQKQHRVCLLHQOQFLTGYSLLMLPLWTSYTF--LSNDQFSTDDF 689  
 QY 662 PCLRADVRVPSESKQCSFYLDKNIHGFYLPASRTSDSQY-DALITSNLVPMVEE 720  
 D 690 SNCLQDLRIPLSPMHKCSYKSTSKLSYGLFTPPRLNRVSRQIYSEALLTSNIVPMYQS 749  
 QY 721 FRKMDYFHSVLLIKHATERNQVNVVSGPIPDYNDGHDFDAPD---ETTKHLANTDVP 777  
 D 750 FQVINYQLHDTVLRVYAQERNVNVVSGVDFDGDYDSEILKQNRVIRSQENLIP 809  
 QY 778 THYFVILSKNKSHTPENCWLDVLPFLPHRTNVVSECEKPEALWVEERFATHA 837  
 D 810 THFFVILSKQLSESPKLCTA-LESSAFLPHRPDNIESTCHGQESAWVEELLALHRA 868  
 QY 838 RYVDVELLTGLDFYDKQVPSSEILQKATYLPTE 871  
 D 869 KYTDVELLTGLSFYQDRQESVSELLRLKTHLP 902

## RESULT 6

ID NPP2 MOUSE STANDARD; PRT; 862 AA.  
 AC Q9RLI6; Q9RLG9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ectonucleotide pyrophosphatase/phosphodiesterase 2 (E-NPP 2)  
 DE (Phosphodiesterase 1/nucleotide pyrophosphatase 2) (Phosphodiesterase  
 DE I alpha) (pD-Ialpha) [includes: Alkaline phosphodiesterase I  
 DE (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].  
 GN ENPP2 OR NPP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RC STRAIN=C57BL/6;  
 RP MEDLINE=20169177; PubMed=10702660;  
 RX Pao J.-H., Matsuda Y., Nakamura H., Sano K.;  
 RA "Assignment of Pdnpp2, the gene encoding phosphodiesterase I/nucleotide  
 RT pyrophosphatase 2, to mouse chromosome 1502.";  
 RL Cytogenet. Cell Genet. 87:172-174(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodríguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides  
 CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated  
 CC oligo-nucleotides.  
 CC -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -!- SIMILARITY: Contains 2 somatomedin-B type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: AF123542; AAD46480.1; -;  
 CC EMBL: BC003264; AA003264.1; -;  
 CC MGD: MGI:1321390; Enpp2.  
 CC InterPro: IPR001604; Endonuclease.  
 CC InterPro: IPR002591; Phosphodiesterase.  
 CC InterPro: IPR001212; Somatomedin-B.  
 CC Pfam: PF01223; Endonuclease; 1.  
 CC Pfam: PF01663; Phosphodiesterase; 1.  
 CC Pfam: PF01033; Somatomedin-B; 2.  
 CC PRINTS: PR00022; SOMATOMEDINB.  
 CC SMART: SM00477; NUC; 1.  
 CC SMART: SM00201; SO; 2.  
 CC PROSITE: PS00524; SOMATOMEDIN\_B; 2.  
 KW Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;  
 KW Multifunctional enzyme.  
 FT DOMAIN 1 11  
 FT TRANSMEM 12 30  
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT CYTOPLASMIC (POTENTIAL)  
 FT (POTENTIAL).  
 FT DOMAIN 31 862  
 FT DOMAIN 54 97  
 FT DOMAIN 98 142  
 FT DOMAIN 144 501  
 FT DOMAIN 597 862  
 FT ACT SITE 209 209  
 FT SITE 126 128  
 FT CARBOHYD 53 53  
 FT CARBOHYD 410 410  
 FT CARBOHYD 524 524  
 FT CARBOHYD 806 806  
 FT CARBOHYD 517 517  
 FT CONFLICT 550 550  
 FT CONFLICT 573 573  
 FT CONFLICT 743 743  
 FT CONFLICT 862 862  
 FT SEQUENCE 862 AA; 98886 MW; C43744DEAA85A355 CRC64;  
 Query Match 43.9%; Score 2111; DB 1; Length 862;  
 Best Local Similarity 44.1%; Pred. No. 2.1e-143;  
 Matches 385; Conservative 171; Mismatches 263; Indels 54; Gaps 10;  
 QY 32 LVMSLGLGLGLG-----LRKLE-----KQSGCRKCFDASFRGLENCR 70  
 D 15 LFTFAIGNVLCGLGFTASRIKRAEDGEPPTVLSDPWNTSGSKGRCFELGVGPPDCR 74





CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
CC EMBL; L35594; AAA64785.1; --
CC EMBL; L46720; AAB00855.1; --
CC EMBL; D45421; BAA08260.1; --
CC EMBL; BC034961; AAH34961.1; --
CC EMBL; HGNC:3357; ENP22.
CC MIM; 601060; --
CC DR GO; GO:0005857; C:integral to plasma membrane; TAS.
CC DR GO; GO:004551; F:nucleotide diphosphatase activity; TAS.
CC DR GO; GO:004528; F:phosphodiesterase I activity; TAS.
CC DR GO; GO:008134; F:transcription factor binding; TAS.
CC DR GO; GO:0006928; P:cell motility; TAS.
CC DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
CC DR GO; GO:0006796; P:phosphate metabolism; TAS.
CC InterPro; IPR001604; Endonuclease.
CC InterPro; IPR002591; Phosphodiester.
CC InterPro; IPR001212; Somatomedin B.
CC Pfam; PF01223; Endonuclease; 1.
CC Pfam; PF01663; Phosphodiester; 1.
CC Pfam; PF01033; Somatomedin B; 2.
CC PRINTS; PR00022; SOMATOMEDINB.
CC SMART; SM00477; NUC; 1.
CC DR SMART; SO; 2.
CC DR PROSITE; PS00524; SOMATOMEDIN B; 2.
CC KW Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW Multifunctional enzyme; Alternative splicing.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 31 863 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 55 98 SOMATOMEDIN-B LIKE 1.
FT DOMAIN 99 143 SOMATOMEDIN-B LIKE 2.
FT DOMAIN 145 502 PHOSPHODIESTERASE.
FT DOMAIN 598 863 NUCLEASE.
FT ACT_SITE 210 210 BY SIMILARITY.
FT SITE 127 129 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 324 324 E -> EESSYGGSPFTPAKPKKAKVAKKRRQRRPAPPKKRR
RKHHRMDHYAAETRODK (in isoform 2).
FT FTID-VSP 006750.
FT N -> S (IN REF. 1).
FT D -> H (IN REF. 3).
FT G -> A (IN REF. 2).
FT Q -> R (IN REF. 1).
FT H -> R (IN REF. 1).
FT P -> L (IN REF. 1).
FT S -> R (IN REF. 3).
FT V -> A (IN REF. 3).
FT Y -> H (IN REF. 3).
FT Y -> H (IN REF. 3).
SQ SEQUENCE 863 AA; 99003 MW; 05B8A2E55C464BC CRC64;
Query Match 43.5%; Score 2091.5; DB 1; Length 863;
Best Local Similarity 43.1%; Pred. No. 5.3e-142;
Matches 381; Conservative 176; Mismatches 267; Indels 59; Gaps 10;
QY 24 ACTVALLAVIMSLGLGLGLGLKLEK-----QGSCKKCFDASF 63
DB 9 SCQLISLFTAVGVNLCLETAHRIKAEWGEGPTVLSDSPWNISGCKRCFELQE 68
QY 64 RGLNCRCDVACRDRGCCWDFEDTCVETRWIMCNKPRCGTRLEASLCSCSDCLQK 123
DB 69 AGPPDCRCDNLKSYTSCCHDFDELCLKTARGWECTKDRCGVRNEENACHCEDCLARG 128
QY 124 DCCADYKSCGGTSLWLENCNCPAQSCQCPGFDLPVILFSDMGFRAEYLTYWDTLMPN 183
DB 129 DCTNTQVUCKGSHWDDCEIKAACPAAGVAPPLIIFSDVGFRAKYKKSKWNP 188
QY 184 INKLKTCGIHSKYRAMYPTKTPFNHYITVTGLPESHGIIIDNNMYDNLNKNFSLSKSKE 243

```

```

DB 189 IEXLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSYDFDPFATHLRGRE 248
QY 244 QNNPAAWHGQPMWLTAHYOGLKAAATYFWPGEVAINGSPGSIYMPYNGSVFEEERISTLL 303
DB 249 KFNHRWGGQPLWITATKGVZAGTFFW-----SWVIPH-----ERRILITL 290
QY 304 KMLDLPKAEPRFYTYWPEEPSSSHAGGPGVAVIKALQVVDHAFGLMGLKQRLNHN 363
DB 291 QMLTLPDHERPSYAFYSEQDPFGHKYGPFGPEMTELRIDKIVQGLMDGLKQLKLR 350
QY 364 CVNIILLADHGMDQVTCNKMXYMTDYFRINPFYVYEGPAPRIR---AHNIPHPDFPSFNS 420
DB 351 CVNVIFVGDHGMEDVTCDETFELSNVLTNVDITLVPCTGLRISKFSNNAKYD-----P 405
QY 421 EELVRLNLSCKDDQHFKPYLTDPDLKRLHYAKNVRIDKHLFVDOOW-----LAVRSKS 474
DB 406 KAIIANLTCKEDQHFKPYLKQHLKRLHYANNRIEDIHLLVRRMHWAKPLDVYKPK 465
QY 475 NTNC-GGGNHGYNNEFRSMEALFLAHGPFSEKTEVEPFENIEVNLNMLCDLLRIQAPANN 533
DB 466 SGKCFEQGDHGFNDKVNNSQTVFCYGTETKTKVPPFENIELYNVWCDLLGLKAPANN 525
QY 534 GTHGSLNHLKVPFYEPESHAEVSKFSVCGFANPLPTESLDQFC-PHLQNSTQLEQVNM 592
DB 526 GTHGSLNHLRTNTRPTMPEEVTNPYPGIMYLSQSDFLGCTCDDKVEPKKLDENKR 585
QY 593 LNLQTEITATVKNLPGRRPRVLOKNDHCLLVHREYVSGFKAMRPMSSYTVPOLG 652
DB 586 LHTK-----GSTERHLLYGRPAVLR-TEYDILVHTDFESGYSEIFLPLMTSYTVSKQA 640
QY 653 DTSPDPTVDDCLRADRVPPSSSKCFYADKNITHGFLYPPASNRTSQSDALITS 712
DB 641 EVSSVPDHLTSCVRPDRVSPFSQNCIAYKXQMSYGLFPFLPYLSSPEAKYDAFLVT 700
QY 713 NLVPMVEEPRKMDYFVHSLIKHATERNVNVYSGPIFDYDGLDHDTEKIKQYVEGS 760
DB 701 NVVPMYPAKRVNYPQVRLVKYASERGVNVLISGPIFDYDGLDHDTEKIKQYVEGS 760
QY 773 DVPIPTHYFVLTSCKNSHTPENCGLDVLPIIHRPTNVNESCPEKPEALWBERF 832
DB 761 SIPVTHYYSITSLDFTQPADKCDGPLSVSSFILPHRPDNESSCSSEDESKWVELM 820
QY 833 TAHIAVRVDELLTGLDFVDQKQVPSBILQKTVLPTFTTI 875
DB 821 KHTARVRDIEHLTSLDFRKTSTERSYPEILLTKYLTHTYSEI 863
RESULT 8
NPP2_RAT
ID_NPP2_RAT STANDARD; PRT; 885 AA.
AC Q646I0;
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 2 (E-NPP 2)
DE (Phosphodiesterase I/nucleotide pyrophosphatase 2) (Phosphodiesterase
DE I alpha) (PD-Ialpha) [includes: Alkaline phosphodiesterase I
DE (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN ENP2 OR NPP2.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Brain;
RC MEDLINE=95050605; PubMed=7961762;
RA Narita M., Goto J., Nakamura H., Sano K.;
RT "Molecular cloning, expression, and localization of a brain-specific
RT phosphodiesterase I/nucleotide pyrophosphatase (pd-Ialpha) from rat
RT brain."
J. Biol. Chem. 269:28235-28242 (1994).

```



RA Fairhead C.A., Faye G., Feldmann H., Fiers W.,  
RA Francinges-Gaillard M.-C., Franco L., Frontali L., Fukuhara H.,  
RA Fuller L.J., Gent M.E., Gigot D., Gilliquet V., Glandsdorff N.,  
RA Goffeau A., Genson M., Grisanti P., Grivell L.A., Haasemann M.,  
RA Hatat D., Hegemann J.H., Herbert C.J., Hilger F., Hohmann S.,  
RA Hollenberg C.P., Huse K., Iborra F., Indge K.J., Isono K., Jackman P.,  
RA Jacq C., Jaquet M., James C.M., Jauniaux J.-C., Jia Y., Jimenez A.,  
RA Kleinhaus U., Kreisel P., Lafranchi G., Lewis C., van der Linden C.G.,  
RA Lucchini G., Lutzenkirchen K., Maat C., Mannheim G., Manzano M.E.,  
RA Martegani E., Mathieu A., Maurer C.T.C., McConnell D., McKee R.A.,  
RA Messenguy F., Mewes H.-W., Molenans F., Montague M.A., Navas L.,  
RA Newlon C.S., Olson M.V., Pallier C., Panzer R.J., Pearson B.M.,  
RA Perea J., Philippsen P., Pierard A., Planta R.J., Plevani P.,  
RA Poetsch B., Pohl F.M., Purnelle B., Ramezani Rad M., Rasmussen S.W.,  
RA Raynal A., Remacha M., Richterich P., Roberts A.B., Rodriguez F.,  
RA Sanz B., Schaaff-Gerstenschlaeger I., Scherens B., Schweitzer B.,  
RA Shu Y., Skala J., Slonimski P.P., Sor F., Soustelle C.,  
RA Spiegelberg R., Staveva L.I., Steensma H.Y., Steiner S., Thierry A.,  
RA Thireos G., Triano L.N., Urrestarazu L.A., Valle G., Vetter I.,  
RA van Vliet-Reedijk J.C., Volckaert G., Vreken P., Warrington J.R.,  
RA von Wettstein D., Wickstead B.L., Wilson C., Wurst H., Xu G.,  
RA Zimmermann F.K., Sgouros J.G.;  
RL "The complete DNA sequence of yeast chromosome III.";  
RL Nature 357:38-46(1992).  
RN [2]  
RN SEQUENCE OF 244-742 FROM N.A.  
RP MEDLINE=92245758; PubMed=1574926;  
RA Bolle P.-A., Gilliquet V., Berben G., Dumont J., Hilger F.;  
RT "The complete sequence of K3B, a 7.9 kb fragment between PGK1 and  
RT CRY1 on chromosome III, reveals the presence of seven open reading  
RT frames.";  
RL Yeast 8:205-213(1992).  
RN [3]  
RN REVISIONS.  
RP Valles G., Volckaerts G.;  
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: TO YEAST YEL016C.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X59720; CAC42978.1; -  
DR PIR; S19437; S19437.  
DR PIR; S27380; S27380.  
DR GerMOnline; 138933; -  
DR SGD; S0000621; YCR026C.  
DR InterPro; IPR002591; Phosphodiester.  
DR Pfam; PF01663; Phosphodiester; 1.  
KW Hypothetical protein; Transmembrane; Glycoprotein.  
FT TRANSMEM 114 135 POTENTIAL.  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 742 AA; 84733 MW; 83BDF5F00D69B09C5 CRC64;  
Query Match 9.0%; Score 434.5; DB 1; Length 742;  
Best Local Similarity 26.5%; Pred. No. 2.3e-23;  
Matches 122; Conservative 84; Mismatches 176; Indels 79; Gaps 18;  
QY 159 PPVILFSGMDGFRABVLYLTWDT-LMPNTNKLKTCG-----IHSKYNRMYPTKTPNHYTIV 213  
Db PLTVISLDGPHPLSLKRWTFPHLDLYELKYDGMNITPTFPWSPFPPTETPNHTLV 227  
QY 214 TGLYPESHGIIDNNYDNNLNKNFSLSKSQNNPAWHG---QPMWLT--AMTQG---LK 265  
Db TGQYPIHHGIVSNVFWFDPLNEEFHGVLD---PRWNNNDTEPIWQTVQSAFDGIPFK 284

QY 266 AATYFWPGSEV-----AINGSEFSTYMPYNGSVPFEEIRISTULKWLDL 308  
Db ANTHWPGSDVNYTKYNEEKLOPHKQPIARERTFFYFDEFNAKEPLSKLSKIETVDM 344  
QY 309 FKA-ERPRFYTWYFEEDSSCHAGGVSAR-----VIKALQVVDHAFGLMEGLKORNL 361  
Db STLNERPQLILGYVNVDAFGHKHGYSESYVEDETEILGEVDTFLKQLVSLQORNL 404  
QY 362 HNCVNIILLADHGMDOYCNK-----MEYMTDYFRINFFYMYEGAPRIRAHNI 411  
Db TSFTNLVIVSDHGMDSIWPVSNVIWEDLLDEKLKDY--VSHAYL-EGPMAISLKD- 459  
QY 412 PHDFFSFNSEIIVRNLSCKRPDQHFKPYLTPDLPKRLHY--AKNVRIDKVLHFDVQWLA 469  
Db SGNINEVYHNLKTSIDEDKYTYVNGNFPKFNVDGNKHWHASIWIVPEPGYAV 514  
QY 470 V-----RSKSNVTCGG-GNHGYNFRSMEAFIAHGPSFKTEVEPPENI 515  
Db MKKQKLVKAKGDKHKNEDNVFTIGSHGYDNNADMRSVFIGWVPYFPQ-GVIEFPQNT 573  
QY 516 EYVNLGCDLLRIQIPAPNNGTHGSLNHLKVPFYEPShAEV 556  
Db EYVNLGCDICGVAEKDRNSDGT--GMLNQLREPOSSEEV 612  
RESULT 10  
YEB6 YEAST  
ID AC YEB6 YEAST STANDARD; PRT; 493 AA.  
DT P39957;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 57.4 kDa protein in PMP2-VAC8 intergenic region.  
GN YEL016C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=97313264; PubMed=9169868;  
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
RA Araujo R., Aviles E., Bero A., Brennan T., Carpenter J., Chen E.,  
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
RL Nature 387:78-81(1997).  
CC -!- SIMILARITY: TO YEAST YCR26C.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U18530; AAB64493.1; -  
DR PIR; S50443; S50443.  
DR GerMOnline; 139020; -  
DR SGD; S0000742; YEL016C.  
DR InterPro; IPR002591; Phosphodiester.  
DR Pfam; PF01663; Phosphodiester; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 493 AA; 57354 MW; 10B67A05C6DEBDF09 CRC64;  
Query Match 8.6%; Score 412.5; DB 1; Length 493;  
Best Local Similarity 28.2%; Pred. No. 5e-22;

Matches 126; Conservative 72; Mismatches 154; Indels 95; Gaps 22;  
QY 150 SQCEGFDLPP-----VILPNDGPRAYL---VTWTLNPNINK 186  
DB 47 SKCFKPDNGERTIANRSNTYFNGTHDFKTLTILSIDGPHRLIDAKYT--PFLYNLEN 104  
QY 187 LKTCG---IHSKYMRYMTKTPPNHYTIVTGLYPESHGIIIDNNYDNLNKNFSLSK 242  
DB 105 LRSFYDNIITAPYWPISFTQFPNHSWVTKYPIEHGIVSNIFWD-----NFTSSEF 159  
QY 243 EONK---PAWQH-QOPW--LTAMYQG-LKAAFYFPGSEVAL--NGSPSPSYMP-----288  
DB 160 RPNLLDARIINSNTADPIWQLLOTESQGEYKVATHMWFGEVYVYDGHGVPFRERMPFYFGK 219  
QY 289 YNGSVPPPEERISTLLKWLDPK-AERPRFYMTYFEBPDGSGHAGG--PVSARVIALQV 345  
DB 220 FNOEKQLDKLAQIFRYIDPQLKORDELVISIYIPNDVSGHSGFDLDELKQLKLGIV 279  
QY 346 DHAFGLMEGLKQRLNLCUNILLADHGMQDTYCNMWEYDY-----FPR-----INFF 396  
DB 280 DGFELDLTEGLQKELNLIKSNVMIVSDHGMNSVYVWVVERVFPADAMSAFISHL 339  
QY 397 YMYEGPAPRIRAHNI PHDFFSFENSEETVRLNLSCKRQDKPKVLTTPDLKRLH-----449  
DB 340 Y-NEGPMWVCLKN-PPD-----KQWICOLIEAQLEKAYGEISRKHFVILKEDF 387  
QY 450 -----YAKNVRIDKVLHVDQQLAVRSKNTNCG-GNHNYN-NFESMEA 494  
DB 388 DPSKYPQYDNKRHYRDVGDVILADEYIAIVKEMGDVPIGIMTHGYNFNNSDKAS 447  
QY 495 IFLAHGSPFKEKTEVEFENIEVYNLM 521  
DB 448 IFICGMPFENNEV-VPFENIEVYNML 473

RESULT 11  
PPD1\_BOVIN STANDARD; PRT; 61 AA.  
AC P15396;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphodiesterase I (EC 3.1.4.1) (5'-nucleotide  
DE phosphodiesterase) (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Intestine; PubMed=2989287;  
RX MEDLINE=85234541; PubMed=2989287;  
RA Culp J.S., Blytt H.J., Hermodson M., Butler L.G.;  
RT "Amino acid sequence of the active site peptide of bovine intestinal  
RT 5'-nucleotide phosphodiesterase and identification of the active  
RT site residue as threonine.";  
RL J. Biol. Chem. 260:8320-8324 (1985).  
CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides  
CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated  
CC oligo-nucleotides.  
DR PIR; A25274; A25274.  
DR InterPro; IPR02591; Phosphodiester.  
DR Pfam; PF01663; Phosphodiester; 1.  
KW Hydrolase.  
FT NON TER 1  
FT ACT SITE 39 39 PHOSPHOTHOEONINE INTERMEDIATE.  
SQ SEQUENCE 61 AA; 6999 MW; 580C8B4807A61C84 CRC64;

Query Match 6.3%; Score 301; DB 1; Length 61;  
Best Local Similarity 86.7%; Pred. No. 2.8e-15;  
Matches 52; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 167 DGPRAEVLWTMDLBNINKLTCGHIHSKYMRYMTKTPPNHYTIVTGLYPESHGIIIDN 226  
DB 1 DGPRAEVLQWSTLVNPNINKLTCGVHVSQYLRAPYPTKTPPNHYTIVTGLYPESHGIIIDN 60  
RESULT 12  
TNAG\_HUMAN STANDARD; PRT; 476 AA.  
AC Q9UW2; Q9UW1; Q9ULZ4;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tubulointerstitial nephritis antigen (TIN-Ag).  
GN TINAG.  
OS Homo sapiens (Human).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=Kidney;  
RX MEDLINE=20119130; PubMed=10652240;  
RA Ikeda M., Takemura T., Hino S., Yoshioka K.;  
RT "Molecular cloning, expression, and chromosomal localization of a  
RT human tubulointerstitial nephritis antigen.";  
RL Biochem. Biophys. Res. Commun. 268:225-230 (2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS TIN1 AND TIN2), TISSUE SPECIFICITY, AND  
RP SUBCELLULAR LOCATION.  
RC TISSUE=Kidney;  
RX MEDLINE=20214190; PubMed=10752525;  
RA Zhou B., Nelson T.R., Kashtan C., Gleason B., Michael A.F., Vlasi M.,  
RA Charonis A.S.;  
RT "Identification of two alternatively spliced forms of human  
RT tubulointerstitial nephritis antigen (TIN-Ag).";  
RL J. Am. Soc. Nephrol. 11:658-668 (2000).  
RN [3]  
RP IDENTIFICATION, AND DISEASE.  
RC TISSUE=Kidney;  
RX MEDLINE=87199690; PubMed=3553704;  
RA Fliger F.D., Wieslander J., Brentjens J.R., Andres G.A.,  
RA Butkowsky R.J.;  
RT "Identification of a target antigen in human anti-tubular basement  
RT membrane nephritis.";  
RL Kidney Int. 31:800-887 (1987).  
RN [4]  
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
RX MEDLINE=92106641; PubMed=1782287;  
RA Butkowsky R.J., Kleppel M.M., Katz A., Michael A.F., Fish A.J.;  
RT "Distribution of tubulointerstitial nephritis antigen and evidence for  
RT multiple forms.";  
RL Kidney Int. 40:838-846 (1991).  
RN [5]  
RP DISEASE.  
RX MEDLINE=93103006; PubMed=1456368;  
RA Katz A., Fish A.J., Santamaria P., Nevins T.E., Kim Y.,  
RA Butkowsky R.J.;  
RT "Role of antibodies to tubulointerstitial nephritis antigen in human  
RT anti-tubular basement membrane nephritis associated with membranous  
RT nephropathy.";  
RL Am. J. Med. 93:691-698 (1992).  
RN [6]  
RP FUNCTION.  
RX MEDLINE=96366847; PubMed=8770961;  
RA Chen Y., Krishnamurti U., Wayner E.A., Michael A.F., Charonis A.S.;  
RT "Receptors in proximal tubular epithelial cells for tubulointerstitial  
RT nephritis antigen.";  
RL Kidney Int. 49:153-157 (1996).  
RN [7]  
RP DEVELOPMENTAL STAGE.  
RX MEDLINE=98305799; PubMed=9643646;  
RA Nelson T.R., Kim Y., Michael A.F., Butkowsky R.J., Charonis A.S.;  
RT "Tubulointerstitial nephritis antigen (TIN-ag) is expressed in

distinct segments of the developing human nephron. ;  
Connect. Tissue Res. 37:53-60(1998).  
-!- FUNCTION: Mediates adhesion of proximal tubule epithelial cells  
via integrins alpha3-beta1 and alphaV-beta3.  
-!- SUBCELLULAR LOCATION: Basement membranes.  
-!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=TIN1;  
IsoId=Q9UJW2-1; Sequence=Displayed;  
Note=Major isoform;  
Name=TIN2;  
IsoId=Q9UJW2-2; Sequence=VSP\_050567, VSP\_050568;  
TISSUE SPECIFICITY: Expressed in the kidney cortex, small  
intestine and cornea.  
-!- DEVELOPMENTAL STAGE: Initially observed in the Bowman's capsule  
during early glomerular capillary loop formation in the kidney. In  
more developmentally mature glomeruli, following transition from  
early to mid-capillary loop stage, expression is higher in the  
proximal tubular basement membrane than in the distal basement  
membrane and Bowman's capsule.  
-!- DISEASE: Antibodies against TINAG are found in sera of patients  
with tubulointerstitial nephritis, a rare autoimmune disorder that  
causes acute and chronic renal injury.  
-!- SIMILARITY: Belongs to peptidase family C1.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; ABC22277; BAA84949.1; -;  
DR EMBL; AF195116; AAF08931.1; -;  
DR EMBL; AF195117; AAF08932.1; -;  
DR PIR; JC7189; JC7189.  
DR HSP; P07688; IQDQ.  
DR MEROPS; C01.973; -;  
DR Genew; HGNC:14599; TINAG.  
DR MIM; 606749; -;  
DR GO; GO:0005604; C:basement membrane; IDA.  
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.  
DR GO; GO:0007155; P:cell adhesion; IDA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR InterPro; IPR009041; PMP\_inhibitor.  
DR InterPro; IPR001212; Somatomedin\_B.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Peptidase\_C1; 1.  
DR SMART; SM00201; SO; 1.  
KW Cell adhesion; Extracellular matrix; Basement membrane; Glycoprotein;  
FT Alternative splicing.  
FT SITE 49 50  
FT GO; GO:0005604; C:basement membrane; IDA.  
FT CLEAVAGE (BY URIN) (POTENTIAL).  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 38 38  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 175 175  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 314 314  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 360 360  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 455 455  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARSPLIC 119 169  
FT Missing (in isoform TIN2).  
FT /FTid=VSP\_050567.  
FT Missing (in isoform TIN2).  
FT /FTid=VSP\_050568.  
FT Missing (in isoform TIN2).  
FT /FTid=VSP\_050569.  
FT FS -> IL (IN REF. 1).  
FT N -> I (IN REF. 2; AAF08932).  
FT P -> L (IN REF. 1).  
FT D -> H (IN REF. 1).  
FT R -> H (IN REF. 1).  
FT L -> R (IN REF. 2; AAF08932).  
FT F -> S (IN REF. 1).  
FT V -> I (IN REF. 1).  
FT SEQUENCE 476 AA; 54646 MW; 32DE88E3083C3077 CRC64;

Query Match 2.7%; Score 130; DB 1; Length 476;  
Best Local Similarity 21.4%; Pred. No. 0.083;  
Matches 56; Conservative 29; Mismatches 69; Indels 108; Gaps 13;  
QY 75 CKDRGDCWDFEDTCVSESTRIMCNKFRGCGTRLEASLCSDDC-LQKDCDCCADYKSV 133  
DB 57 CRNFG-CCEDRDGCV--TEFYAN-----ALCYCDKFCDENSDCCPDYKSF 102  
QY 134 QGETSW-----LEENCD-----TAQSQCPGFDL--PPVI- 162  
DB 103 REEKWPPHTOPWYPEGCFKDGQHYEGSVIKENCNCTCGQGWKCSQHVCLVRPELIE 162  
QY 163 -----LFSMDGFRAEVLYTWTLMFNINKTKTGIIHSKYMRMYPT 203  
DB 163 QVNGDYGVMTAQNVSQFWGMTLEDGFK-----FRIGTLPPSPMLLS-----MNMETASLPA 213  
QY 204 KTFNHYTIVTGLYPE-SHGIIID-----NNMYDYNLKNKFS 238  
DB 214 TTDLPFEFFVASYKWPQWTHGFLDKQNCASWAFSTASVAADRIAIOSKGRYTANLSPQNL 273  
QY 239 LSSKBQN-----NPAWH 251  
DB 274 ISCAKNRHCNGSGSIDRAWWY 295  
-----  
RESULT 13  
P111 HUMAN  
ID P111 HUMAN STANDARD; PRT; 369 AA.  
AC P21128;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Placental protein 11 precursor (EC 3.4.21.-) (P111).  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-29.  
RC TISSUE=Placenta;  
RX MEDLINE=90274869; PubMed=2350438;  
RA Grunemann U., Roemisch J., Siebold B., Bohn H., Amann E.;  
RT "Cloning and expression of a cDNA encoding human placental protein  
11, a putative serine protease with diagnostic significance as a  
tumor marker.";  
RL DNA Cell Biol. 9:243-250(1990).  
RN [2]  
RP SOMATOMEDIN-B TYPE DOMAIN.  
RX MEDLINE=91248172; PubMed=1710108;  
RA Jenne D.;  
RT "Homology of placental protein 11 and pea seed albumin 2 with  
vitronectin.";  
RL Biochem. Biophys. Res. Commun. 176:1000-1006(1991).  
CC -!- FUNCTION: Probable serine protease.  
CC -!- MISCELLANEOUS: PLACENTAL PROTEIN 11 IS A PLACENTAL-SPECIFIC  
CC PROTEIN BUT IS ALSO ASSOCIATED WITH VARIOUS MALIGNANT NEOPLASMS.  
CC -!- SIMILARITY: Contains 1 somatomedin-B type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M32402; AAA36464.1; -;  
DR EMBL; M36109; AAA36465.1; -;  
DR PIR; A34614; A34614.  
DR MIM; 606720; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
DR



InterPro: IPR001212; Somatomedin B.  
Pfam: PF01033; Somatomedin B; 1.  
PRINTS: PR00022; SOMATOMEDINB.  
SMART: SM00201; SO; 1; SOMATOMEDIN B; 1.  
PROSITE: PS00524; SOMATOMEDIN B; 1.  
Hydrolase; Serine protease; Signal; Polymorphism.  
SIGNAL 1 18 PLACENTAL PROTEIN 11.  
CHAIN 19 369 SOMATOMEDIN-B LIKE.  
DOMAIN 45 89  
VARIANT 31 31  
VARIANT 31 31  
VARIANT 31 31  
VARIANT 31 31  
SEQUENCE 369 AA; 42121 MW; F5935AE12D7E924C CRC64;  
Query Match 2.7%; Score 129; DB 1; Length 369;  
Best Local Similarity 19.6%; Pred. No. 0.069;  
Matches 71; Conservative 48; Mismatches 104; Indels 140; Gaps 17;  
QY 53 SCRKCFDASPRGLNCRDCAKDRGDCWDFEDTCVSTRIMCNKFRGCTRLASL 112  
DB 48 SQGRCYEA-FDKHQCHCNARCQFGNCKDFESLCSDE----- 87  
QY 113 CSCSDCLCKDCCADYKSVCGETSLE-----ENC-----DTAQ-SOCPEGFDLP 159  
DB 88 VSHSDAITKBEIQISEKIYRADINKAQEDIVLNSQNCISPSSTRNQVRCPKP----- 143  
QY 160 PVILFSMDGFRAEYLYTMDLMPNINKLTCGIHSKYRMYPTKTFPHNYTIVTGLYPE 219  
DB 144 ---LFT---YVNEKLFSEKTYAAFINLL-----NNYQRTAG----- 173  
QY 220 SGHITDNNMYDNLNKNFSLSKSQNNPAAWQGMWLTAMYQGLKAATYFPGSEVAIN 279  
DB 174 -HG-----EHFSAQELABQDAFL-----REIMKTAVMKELY----- 203  
QY 280 GSFPSTYMPYNGSVPEERISTLLKWLDPKAERPRFTYMTYFEPDSSGHAGGVSARVI 339  
DB 204 -SFLHQNTYSGEQEFVDLKNM--WFGL-----YSRNEEGDSSG----- 241  
QY 340 KALQVVDHAFGLMEGLKORNHNCVNIILLADHGMQDQYCNKMEYMTDYFPRINFFMY 399  
DB 242 -----FEHFVSGEVKKGKVTGFHNWIRF-----YLBEKGLVDYYSHI-----Y 280  
QY 400 EGP 402  
DB 281 DGP 283  
RESULT 14  
ID\_YP73\_YERPE STANDARD; PRT; 2004 AA.  
AC Q8ZD2; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0192 protein YP02573/Y1143 precursor.  
GN YP02573 OR Y1143.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moulie S., Oyston P.C.F., Quail M.A., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague."

Nature 413:523-527(2001).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=KIMS / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Dang W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM."  
RL J. Bacteriol. 184:4601-4611(2002).  
CC -I- SIMILARITY: Belongs to the UPF0192 family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AJ414152; CAC91375.1; -;  
CC EMBL; AE013717; AM84721.1; -;  
CC PIR; AC0314; AC0314.  
KW Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 2004 HYPOTHETICAL PROTEIN YP02573.  
SQ SEQUENCE 2004 AA; 224130 MW; EA9F8A1F7452672C CRC64;  
Query Match 2.6%; Score 126; DB 1; Length 2004;  
Best Local Similarity 19.7%; Pred. No. 1.1;  
Matches 131; Conservative 89; Mismatches 258; Indels 186; Gaps 31;  
QY 236 NFSLSKSEQNNPAAWQGMWLTAMYQGLKAATYFPGSEVAINSGSP-SIYMPYNGSV 294  
DB 283 NFSVYVYDEKLNIAHISQ--LMPDRG-----GSVHLSINKGNASVA 324  
QY 295 FEERISTLLKWLDPKAERPRFTYMTYFEPDSSGHAGGVSARVIKALQVVDHAFG 350  
DB 325 ATPITQANQKVSV-----PNLYSL-----AVNS-NATIVES-----DNNNGERAL 365  
QY 351 --MLMEGLKORNHNCVNIILLADHGMQDQYCNKMEYMTDYFP----- 391  
DB 366 IIAISDAVKKEIKNAVKAWLLPOHNFQAKESAKTS--TDYFPMDDDDIDLLQOSTPL 423  
QY 392 --RIN-----FFYMEGPAPRIAHNIPHDFFSNSEEIVRNLSCKRPQHFKPYL 440  
DB 424 ALTNEAEQEQYQPIFSFKEDAPSVRTLLIEVNNLSLTSVGGYKM-----PEKIYQIVR 475  
QY 441 TPDLPKRLHYAKNVRIDKVLHFDVQOMLAVRSKNTNGCGNGHGYNNFESMEALFLAHG 500  
DB 476 VPDYPTKLRFSQGSLS--LSMQGDKQ--ISVAARNMT-----GMKLDIKRVIPSQLQHI 525  
QY 501 PSFKTEVEFPFENIEVYNLWCDLLRIQAPNNGTHGSLNH----- 541  
DB 526 VSFKSSEYSSAHFNRLSDEYTFHFQYOTALLNNDNPEINQGVDSLSRYLANNSARRGV 585  
QY 542 -LLKVPFYE-----SHAE-----VSFSGVGFANPLPTESLDCFCPHLQNST 584  
DB 586 FLITLSAWDPEKRDNQHQSEEDYDDEQEWGDSFRVVTIDLGITIKSQSD-----RSRDV 640  
QY 585 QLEQVQNMNLTOEITATVKNLPPGPRVLQKNVDHCLLYHREYVSGFGKAMPMWS 644  
DB 641 FVQSIHSGLPAAADAKVSVVAKNGV-----LLSQITDSKGVHVPFALDAF-KNERQPV-- 692  
QY 645 SVTVPOLGDTSPPTVP-----DCLRADV--RVPPSESQKCSFYL-ADKNITHGFLYPP 696  
DB 693 MFLVKEGDSVFLPRTATYDRNLDSRFDIDGEEFSPDRTLSYLSYFSDRGV----- 744  
QY 697 ASNRISDSQYDALITSNL-----VPMYEFKKNWYFHSVLLIKATERNGVNVVS- 747  
DB 745 --YRCPDRENIGLITRTANWATDGVPLRAEIRDPRTLMSTLPI--TLSSSGFNEUSY 800



```

QY 748 -----GPIFDYN-YDGHFADPDEITKILANTDVPi-----PTHFVVLTSCKNKSHTPENC 797
Db 801 TTGENSTGEWVYLYLVGKNNETSMLLGHTTVNVEFEDRLKVL-----QLTPERQ 854

QY 798 PGWL 801
Db 855 QGWV 858

RESULT 15
LRPI_CHICK
ID LRPI_CHICK STANDARD; PRT: 4543 AA.
AC P98157;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Low-density lipoprotein receptor-related protein 1 precursor (LRP)
DE (Alpha-2-macroglobulin receptor) (A2MR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=White leghorn; TISSUE=Liver, and Ovary;
RA MEDLINE=94103212; PubMed=7506255;
RX Nimpf J., Scifani S., Bilous P.T., Schneider W.J.;
RT "The somatic cell-specific low density lipoprotein receptor-related
RT protein of the chicken. Close kinship to mammalian low density
RT lipoprotein receptor gene family members.";
RL J. Biol. Chem. 269:212-219(1994).
CC -!- FUNCTION: Involved in the plasma clearance of chylomicron remnants
CC and activated alpha 2-macroglobulin, as well as the local
CC metabolism of complexes between plasminogen activators and their
CC endogenous inhibitors. Binds vitellogenin, calcium and alpha 2-
CC macroglobulin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC . Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P98157-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P98157-2; Sequence=VSP_004312;
CC -!- TISSUE SPECIFICITY: Somatic.
CC -!- PTM: Cleaved into a 85 kDa membrane-spanning subunit (LRP-85) and
CC a 515 kDa large extracellular domain (LRP-515) that remains non-
CC covalently associated.
CC -!- SIMILARITY: Contains 22 EGF-like domains.
CC -!- SIMILARITY: Contains 31 LDL-receptor class A domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74904; CAA52870.1; -
CC PIR; A53102; A53102.
CC PDB; 1LPX; 29-DEC-99.
CC InterPro; IPR000152; Asx hydroxyl_S.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00057; EGF_14.
CC Pfam; PF00058; ldl_recept_b; 33.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 3.

```

```

DR SMART; SM00192; LDLg; 31.
DR SMART; SM00135; LY; 34.
DR PROSITE; PS00010; ASX HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 7.
DR PROSITE; PS00026; EGF_3; 8.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 31.
KW Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein; Signal;
KW Calcium-binding; EGF-like domain; Coated pits; Alternative splicing;
KW 3D-structure.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 4543 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT DOMAIN 22 4419 PROTEIN 1.
FT TRANSMEM 4420 4443 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 4444 4543 POTENTIAL.
FT DOMAIN 27 68 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 72 112 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 113 151 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 152 191 EGF-LIKE 1.
FT DOMAIN 476 522 EGF-LIKE 2.
FT DOMAIN 801 841 EGF-LIKE 3.
FT DOMAIN 850 890 EGF-LIKE 4.
FT DOMAIN 891 931 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 932 971 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 972 1011 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 1011 1051 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 1058 1097 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 1100 1140 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 1141 1180 LDL-RECEPTOR CLASS A 9.
FT DOMAIN 1181 1220 LDL-RECEPTOR CLASS A 10.
FT DOMAIN 1221 1260 EGF-LIKE 5.
FT DOMAIN 1534 1577 EGF-LIKE 6.
FT DOMAIN 1842 1883 EGF-LIKE 7.
FT DOMAIN 2151 2191 EGF-LIKE 8.
FT DOMAIN 2472 2512 EGF-LIKE 9.
FT DOMAIN 2516 2557 EGF-LIKE 10.
FT DOMAIN 2558 2596 LDL-RECEPTOR CLASS A 11.
FT DOMAIN 2597 2635 LDL-RECEPTOR CLASS A 12.
FT DOMAIN 2636 2684 LDL-RECEPTOR CLASS A 13.
FT DOMAIN 2688 2730 LDL-RECEPTOR CLASS A 14.
FT DOMAIN 2730 2769 LDL-RECEPTOR CLASS A 15.
FT DOMAIN 2770 2812 LDL-RECEPTOR CLASS A 16.
FT DOMAIN 2814 2853 LDL-RECEPTOR CLASS A 17.
FT DOMAIN 2854 2897 LDL-RECEPTOR CLASS A 18.
FT DOMAIN 2900 2938 LDL-RECEPTOR CLASS A 19.
FT DOMAIN 2939 2978 LDL-RECEPTOR CLASS A 20.
FT DOMAIN 2979 3019 EGF-LIKE 11.
FT DOMAIN 3019 3058 EGF-LIKE 12.
FT DOMAIN 3287 3328 EGF-LIKE 13.
FT DOMAIN 3329 3368 LDL-RECEPTOR CLASS A 21.
FT DOMAIN 3369 3407 LDL-RECEPTOR CLASS A 22.
FT DOMAIN 3408 3447 LDL-RECEPTOR CLASS A 23.
FT DOMAIN 3448 3488 LDL-RECEPTOR CLASS A 24.
FT DOMAIN 3489 3530 LDL-RECEPTOR CLASS A 25.
FT DOMAIN 3531 3569 LDL-RECEPTOR CLASS A 26.
FT DOMAIN 3570 3608 LDL-RECEPTOR CLASS A 27.
FT DOMAIN 3608 3646 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3649 3689 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3690 3730 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3736 3776 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3779 3821 EGF-LIKE 14.
FT DOMAIN 3822 3859 EGF-LIKE 15.
FT DOMAIN 4146 4182 EGF-LIKE 16.
FT DOMAIN 4195 4231 EGF-LIKE 17.
FT DOMAIN 4231 4267 EGF-LIKE 18.
FT DOMAIN 4267 4303 EGF-LIKE 19.
FT DOMAIN 4303 4339 EGF-LIKE 20.
FT DOMAIN 4339 4374 EGF-LIKE 21.
FT DOMAIN 4374 4409 EGF-LIKE 22.
FT SITE 3939 3942 RECOGNITION SITE FOR PROTEOLYTICAL
FT PROCESSING (POTENTIAL).

```



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 12:45:12 ; Search time 64 Seconds  
(without alignments)  
3862.954 Million cell updates/sec

Title: US-10-005-480A-743  
Perfect score: 4804  
Sequence: 1 MESTLTLATQPVKXNLTLLK.....QPVSEILQLKYLPTFTTTI 875

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A Geneseq 29Jan04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4804	100.0	875	7 ADE56099	Human Pro
2	4804	100.0	875	7 ADE56103	Human Pro
3	4804	100.0	875	7 ADE55891	Human 161
4	4804	100.0	875	7 ADE55904	Human 161
5	4804	100.0	875	7 ADE55969	Human 161
6	4804	100.0	875	7 ADE55899	Human 161
7	4804	100.0	875	7 ADE55983	Human 161
8	4804	100.0	875	7 ADE55982	Human 161
9	4804	100.0	875	7 ADE55909	Human 161
10	4804	100.0	875	7 ADE55984	Human 161
11	4804	100.0	875	7 ADE55991	Human 161
12	4801	99.9	875	7 ADE55910	Human 161
13	4801	99.9	875	7 ADE55905	Human 161
14	4801	99.9	875	7 ADE55893	Human 161
15	4798	99.9	875	7 ADE55897	Human 161
16	4798	99.9	875	7 ADE55907	Human 161
17	4797	99.9	875	7 ADE55906	Human 161
18	4797	99.9	875	7 ADE55895	Human 161
19	4796	99.8	875	7 ADE55901	Human 161
20	4795	99.8	875	6 ABG74664	Human ecc
21	4743	98.7	885	6 ABG10289	Novel hum
22	4642	96.6	841	7 ADE55989	Human 161
23	4642	96.6	841	7 ADE55903	Human 161
24	4634	96.5	841	7 ADE55990	Human 161
25	4634	96.5	841	7 ADE55988	Human 161

26	4634	96.5	841	7 ADE56098	Ade56098 Human 161
27	3953.5	82.3	876	7 ADE56097	Ade56097 Rat Prote
28	3953.5	82.3	876	7 ADE56101	Ade56101 Rat Prote
29	2498.5	52.0	873	3 AAY83620	Aay83620 Human PC-
30	2498.5	52.0	925	2 AAR79148	Aar79148 Human ins
31	2498.5	52.0	925	2 AAY39355	Aay39355 Insulin r
32	2498.5	52.0	925	3 AAB00195	Aab00195 Breast ca
33	2498.5	52.0	925	7 ADD46549	Add46549 Human Pro
34	2494.5	51.9	873	3 AAY83621	Aay83621 Variant h
35	2265	47.1	945	4 AAM25587	Aam25587 Human pro
36	2122.5	44.2	858	4 AAY71998	Aay71998 Rat autot
37	2120	44.1	862	5 ABG32516	Abg32516 Rat lysop
38	2119.5	44.1	858	4 AAY71986	Aay71986 Rat autot
39	2119.5	44.1	858	4 AAY71995	Aay71995 Rat autot
40	2118.5	44.1	858	4 AAY71997	Aay71997 Rat autot
41	2118.5	44.1	858	4 AAY71996	Aay71996 Rat autot
42	2116.5	44.1	858	4 AAY71988	Aay71988 Human ter
43	2094.5	43.6	863	7 ADC10214	Adc10214 Human NOV
44	2094.5	43.6	863	7 ADC10214	Adc10214 Human NOV
45	2091.5	43.5	870	7 ADC10224	Adc10224 Human NOV

ALIGNMENTS

RESULT 1  
ADE56099  
ID ADE56099 standard; protein; 875 AA.  
XX  
AC ADE56099;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein AAC51813, SEQ ID NO 1940.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
FN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GSHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; AAC51813.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPVKNTLAKYKACIALVIMSLGILGLGRLKLEKQSCRRKKCFD 60  
 DB 1 MESTLTATEQPVKNTLAKYKACIALVIMSLGILGLGRLKLEKQSCRRKKCFD 60  
 QY 61 ASFRGLENCRCVACKDRGCCWEDFEDTCTVESTRIWMCNFKCGTRLEASLCSGDDCL 120  
 DB 61 ASFRGLENCRCVACKDRGCCWEDFEDTCTVESTRIWMCNFKCGTRLEASLCSGDDCL 120  
 QY 121 QKDCCADYKSCQGETSLENCUTAOQSCPEGFDPVILFSDMGFRAEYLYTWDTL 180  
 DB 121 QKDCCADYKSCQGETSLENCUTAOQSCPEGFDPVILFSDMGFRAEYLYTWDTL 180  
 QY 181 MPNINKLTCGISHKVMRAMYPTKTPHNYTIVTGLYPESHGIIIDNNMYDNLKNFSL 240  
 DB 181 MPNINKLTCGISHKVMRAMYPTKTPHNYTIVTGLYPESHGIIIDNNMYDNLKNFSL 240  
 QY 241 SKEQNPAAWHQOPWMLTAMYQGLKAAATYFPGSEVAINGSPPSIYMPYNGSVPFERIS 300  
 DB 241 SKEQNPAAWHQOPWMLTAMYQGLKAAATYFPGSEVAINGSPPSIYMPYNGSVPFERIS 300  
 QY 301 TLLKWLDPKAEPRPRFYTWYFEEPDSSGHAGSPVSARVICALQVVDHAFGLMGLKORN 360  
 DB 301 TLLKWLDPKAEPRPRFYTWYFEEPDSSGHAGSPVSARVICALQVVDHAFGLMGLKORN 360  
 QY 361 LHCNVNIIILLADHGMDOQYCNKMEYTDYFPRINFFYMEGPAIRAHNIPHPDFFS 420  
 DB 361 LHCNVNIIILLADHGMDOQYCNKMEYTDYFPRINFFYMEGPAIRAHNIPHPDFFS 420  
 QY 421 EIVNLSCKRQDOHKEVLTDPKRLHYAKNVRIDKVLHFDQOOLAVRSKNTCCG 480  
 DB 421 EIVNLSCKRQDOHKEVLTDPKRLHYAKNVRIDKVLHFDQOOLAVRSKNTCCG 480  
 QY 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVNLMDCLLRIOAPNNGTHGSLN 540  
 DB 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVNLMDCLLRIOAPNNGTHGSLN 540  
 QY 541 HLLKVPFEPHSAEVSFSGFANPLPTESLDCFCPLHNSQLSOVNOMLNLTQEEI 600  
 DB 541 HLLKVPFEPHSAEVSFSGFANPLPTESLDCFCPLHNSQLSOVNOMLNLTQEEI 600  
 QY 601 TATVKVNLFPGRPRVLQKNVDRCLLYHREYVSGFKAMRPMWSSYTVPOLGDTSPUPPT 660  
 DB 601 TATVKVNLFPGRPRVLQKNVDRCLLYHREYVSGFKAMRPMWSSYTVPOLGDTSPUPPT 660  
 QY 661 VPDCLRADVRVPPESSQKCSYLADKNTHTGFLVPPASNTSDSQYDALITSNLVPMYEE 720  
 DB 661 VPDCLRADVRVPPESSQKCSYLADKNTHTGFLVPPASNTSDSQYDALITSNLVPMYEE 720

QY 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHDAPDEITGHANTDVPPIPHY 780  
 DB 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHDAPDEITGHANTDVPPIPHY 780  
 QY 781 FVLTSCNKSHTPENCPCGMLDVLPIIHRPTNVSCPEGKPEALWVERFTAHRVR 840  
 DB 781 FVLTSCNKSHTPENCPCGMLDVLPIIHRPTNVSCPEGKPEALWVERFTAHRVR 840  
 QY 841 DVELLTGLDFYQDKVQVPSILQIKTYLPTFETI 875  
 DB 841 DVELLTGLDFYQDKVQVPSILQIKTYLPTFETI 875

# RESULT 2

AD56103

ID AD56103 standard; protein; 875 AA.

XX AC AD56103;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein AAC51813, SEQ ID NO 1944.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX FN WC2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; AAC51813.

XX PT New composition comprising two or more isolated polypeptides, useful for

XX PS preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction



Db 241 SKQNNPAWHGQPMMLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVPPERIS 300  
Qy 301 TLLKWLDPKAERPRFYTYMFEPPDSSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
Db 301 TLLKWLDPKAERPRFYTYMFEPPDSSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
Qy 361 LHCNVNILLADGMDQTYCNKMEYMTDYPFRINFFYMEGAPAPRAHNI PHDFSFNS 420  
Db 361 LHCNVNILLADGMDQTYCNKMEYMTDYPFRINFFYMEGAPAPRAHNI PHDFSFNS 420  
Qy 421 EEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCGG 480  
Db 421 EEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCGG 480  
Qy 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Db 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Qy 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTOLEQVNMNLNQBEI 600  
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTOLEQVNMNLNQBEI 600  
Qy 601 TATVKVNLPEGRPRVLOKNDHCLLYHREYVSGFGKAMRMPMWSYTVPOLGDTSPPLPT 660  
Db 601 TATVKVNLPEGRPRVLOKNDHCLLYHREYVSGFGKAMRMPMWSYTVPOLGDTSPPLPT 660  
Qy 661 VPCLRADVRVPSESKCSFYLDADKNITGHFLYPPASNRTSDSOYDALITSNLVPMEY 720  
Db 661 VPCLRADVRVPSESKCSFYLDADKNITGHFLYPPASNRTSDSOYDALITSNLVPMEY 720  
Qy 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPPIPTHY 780  
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPPIPTHY 780  
Qy 781 FVVLTSCKNKSHPENCGLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIAVR 840  
Db 781 FVVLTSCKNKSHPENCGLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIAVR 840  
Qy 841 DVELLTGLDFYQDKQVPVSEIILQLKTYLPTFTTI 875  
Db 841 DVELLTGLDFYQDKQVPVSEIILQLKTYLPTFTTI 875

## RESULT 4

AD65904  
ID AD65904 standard; protein; 875 AA.  
XX AC AD65904;  
XX AC AD65904;  
DT 29-JAN-2004 (first entry)  
XX DE Human 161P2F10B protein variant 1 #2.  
XX KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.  
XX OS Homo sapiens.  
XX FN WO2003040340-A2.  
XX PD 15-MAY-2003.  
XX PF 07-NOV-2002; 2002WO-05036002.  
XX PR 07-NOV-2001; 2001US-00005480.  
XX PR 31-JAN-2002; 2002US-00062109.  
XX PA (AGEN-) AGENSYS INC.  
XX PI Jakobovits A, Raitano AB, Paris M, Hubert RS, Ge W, Morrison KJM;  
XX PI Morrison RK, Challita-Eld PM;  
XX WPI; 2003-441560/41.  
DR N-PSDB; AD65991.

XX A composition for diagnosing, preventing and treating cancer (e.g.  
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
PT and polypeptides.  
XX Example 1; SEQ ID NO 16; 135pp; English.  
XX This invention relates to a novel composition which comprises a substance  
CC that modulates the status of a novel protein (161P2F10B) and its variants  
CC having a sequence of 875 amino acids provided in the specification. The  
CC protein of the invention is over-expressed in certain cancers. The  
CC compounds of the invention may have cytostatic activity and the sequence  
CC of the 161P2F10B protein, and the gene which encodes it, may be useful  
CC for gene therapy or the development of a vaccine. The composition and  
CC methods of the invention are useful in diagnosing, preventing and  
CC treating cancer. The present sequence is the amino acid sequence of the  
CC human 161P2F10B variant 1 protein of the invention.  
XX Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MESTLTTLTATEQPVKQNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKCCF 60  
Db 1 MESTLTTLTATEQPVKQNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKCCF 60  
Qy 61 ASRPLENCDCVACKDRGDCWDFTCVETRIWNCNKRFCGETRLEASLCSDDCL 120  
Db 61 ASRPLENCDCVACKDRGDCWDFTCVETRIWNCNKRFCGETRLEASLCSDDCL 120  
Qy 121 QKQCCADYKSVQGETSWLEENCDAQAQSCQEGFDLPVILFMSDGFRAEYLYTWDTL 180  
Db 121 QKQCCADYKSVQGETSWLEENCDAQAQSCQEGFDLPVILFMSDGFRAEYLYTWDTL 180  
Qy 181 MPNINKLKTGCIHSKYRAMYPTKTPFNHYITVGLYPESHGIIIDNMVDVNLKQFSL 240  
Db 181 MPNINKLKTGCIHSKYRAMYPTKTPFNHYITVGLYPESHGIIIDNMVDVNLKQFSL 240  
Qy 241 SKQNNPAWHGQPMMLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVPPERIS 300  
Db 241 SKQNNPAWHGQPMMLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVPPERIS 300  
Qy 301 TLLKWLDPKAERPRFYTYMFEPPDSSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
Db 301 TLLKWLDPKAERPRFYTYMFEPPDSSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
Qy 361 LHCNVNILLADGMDQTYCNKMEYMTDYPFRINFFYMEGAPAPRAHNI PHDFSFNS 420  
Db 361 LHCNVNILLADGMDQTYCNKMEYMTDYPFRINFFYMEGAPAPRAHNI PHDFSFNS 420  
Qy 421 EEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCG 480  
Db 421 EEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCG 480  
Qy 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Db 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Qy 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTOLEQVNMNLNQBEI 600  
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTOLEQVNMNLNQBEI 600  
Qy 601 TATVKVNLPEGRPRVLOKNDHCLLYHREYVSGFGKAMRMPMWSYTVPOLGDTSPPLPT 660  
Db 601 TATVKVNLPEGRPRVLOKNDHCLLYHREYVSGFGKAMRMPMWSYTVPOLGDTSPPLPT 660  
Qy 661 VPCLRADVRVPSESKCSFYLDADKNITGHFLYPPASNRTSDSOYDALITSNLVPMEY 720  
Db 661 VPCLRADVRVPSESKCSFYLDADKNITGHFLYPPASNRTSDSOYDALITSNLVPMEY 720  
Qy 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPPIPTHY 780



```
|||||
721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHFDADPEITKHLANTDVPITPHY 780
781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCPGKPEALWVEERTAHIAVR 840
781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCPGKPEALWVEERTAHIAVR 840
841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875
841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875

RESULT 5
ADE65899 standard; protein; 875 AA.
XX
AC ADE65899;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 161P2F10B protein variant 1 #4.
XX
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS Homo sapiens.
XX
PN WO2003040340-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US036002.
XX
PR 07-NOV-2001; 2001US-00005480.
XX
PR 31-JAN-2002; 2002US-00062109.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Raitano AB, Farris M, Hubert RS, Ge W, Morrison KJM;
PI Morrison RK, Challita-Eid PM;
XX
DR WPI; 2003-441560/41.
XX
XX
XX A composition for diagnosing, preventing and treating cancer (e.g.
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT and polypeptides.
XX
XX Example 65; SEQ ID NO 81; 135pp; English.
XX
CC This invention relates to a novel composition which comprises a substance
CC that modulates the status of a novel protein (161P2F10B) and its variants
CC having a sequence of 875 amino acids provided in the specification. The
CC protein of the invention is over-expressed in certain cancers. The
CC compounds of the invention may have cytostatic activity and the sequence
CC of the 161P2F10B protein, and the gene which encodes it, may be useful
CC for gene therapy or the development of a vaccine. The composition and
CC methods of the invention are useful in diagnosing, preventing and
CC treating cancer. The present sequence is the amino acid sequence of the
CC human 161P2F10B variant 1 protein of the invention.
XX
SQ Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTLATEQPVKNTKKYKIACIVLLALVIMSLGGLGRLKLEKQSGCRKKCFD 60
Db 1 MESTLTLATEQPVKNTKKYKIACIVLLALVIMSLGGLGRLKLEKQSGCRKKCFD 60
QY 61 ASFRGLENCRCDVACDRGDCWDFEDTCVETRIWMCNKRFCGETRLEASLCSDDCL 120
Db 61 ASFRGLENCRCDVACDRGDCWDFEDTCVETRIWMCNKRFCGETRLEASLCSDDCL 120
```

```
QY 121 QKDCCADYKSVQCGETSWELENCDTAQSCCEGFDLPVILLFSMDGFAEYLYTWDTL 180
Db 121 QKDCCADYKSVQCGETSWELENCDTAQSCCEGFDLPVILLFSMDGFAEYLYTWDTL 180
QY 181 MPNINKLKTGCIHSHKYRMYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240
Db 181 MPNINKLKTGCIHSHKYRMYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240
QY 241 SKEQNPAWHGQPMWLTYMGLKAATYFWPGEVAINGSFSPSIYMPYNGSVPEERIS 300
Db 241 SKEQNPAWHGQPMWLTYMGLKAATYFWPGEVAINGSFSPSIYMPYNGSVPEERIS 300
QY 301 TLLKWLDPKAERPRFYTMVFEEDPSSGHAGGPPVSARVVKALQVVDHAFGLMGLKORN 360
Db 301 TLLKWLDPKAERPRFYTMVFEEDPSSGHAGGPPVSARVVKALQVVDHAFGLMGLKORN 360
QY 361 LHNCVNIILLADHGMDOYCNKMEYMTDYPRINFFYNYEGPAPRIRAHNIPHDFFSFNS 420
Db 361 LHNCVNIILLADHGMDOYCNKMEYMTDYPRINFFYNYEGPAPRIRAHNIPHDFFSFNS 420
QY 421 BEIVRNLSCKRPDOHFKPYLTPDLPKELHYAKNVRIDKVLHFVDOQWLAVRSKNTKCG 480
Db 421 BEIVRNLSCKRPDOHFKPYLTPDLPKELHYAKNVRIDKVLHFVDOQWLAVRSKNTKCG 480
QY 481 GNHCYNNEFRSMEAFILAHGSPFKEKTEVEPFENIEVYNLMCDLLRQPAPNNGTHGSLN 540
Db 481 GNHCYNNEFRSMEAFILAHGSPFKEKTEVEPFENIEVYNLMCDLLRQPAPNNGTHGSLN 540
QY 541 HLLKVPYFESHABEVSKFSVCGFANPLPTESLDCTCFHLQNSTQLEQVQMLNLTQEEI 600
Db 541 HLLKVPYFESHABEVSKFSVCGFANPLPTESLDCTCFHLQNSTQLEQVQMLNLTQEEI 600
QY 601 TATVKVNLPGCRPRVLQKNDHCLLYHREYVSGFGKAMRMPKSSYTVPOLGDTSPLPPT 660
Db 601 TATVKVNLPGCRPRVLQKNDHCLLYHREYVSGFGKAMRMPKSSYTVPOLGDTSPLPPT 660
QY 661 VPDCLRADRVPPSESQKCSFYLDXKNITGHFLYPPASNETSDQYDALITSNLVPMYEE 720
Db 661 VPDCLRADRVPPSESQKCSFYLDXKNITGHFLYPPASNETSDQYDALITSNLVPMYEE 720
QY 721 PRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHFDADPEITKHLANTDVPITPHY 780
Db 721 PRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHFDADPEITKHLANTDVPITPHY 780
QY 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCPGKPEALWVEERTAHIAVR 840
Db 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCPGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875
Db 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875

RESULT 6
ADE65899
ID ADE65899 standard; protein; 875 AA.
AC ADE65899;
XX
XX 29-JAN-2004 (first entry)
XX
DE Human 161P2F10B protein variant 5.
XX
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS Homo sapiens.
XX
PN WO2003040340-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US036002.
XX
```

PR 07-NOV-2001; 2001US-00005480.  
PR 31-JAN-2002; 2002US-00062109.  
XX (AGEN-) AGENSYS INC.  
XX  
PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KM;  
PI Morrison RK, Challita-Eid PM;  
XX N-PSDB; ADE65983.  
DR WPI: 2003-441560/41.  
XX N-PSDB; ADE65983.  
XX  
PT A composition for diagnosing, preventing and treating cancer (e.g.  
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
PT and polypeptides.  
XX  
PS Claim 1; SEQ ID NO 11; 135pp; English.  
XX  
CC This invention relates to a novel composition which comprises a substance  
CC that modulates the status of a novel protein (161P2F10B) and its variants  
CC having a sequence of 875 amino acids provided in the specification. The  
CC protein of the invention is over-expressed in certain cancers. The  
CC compounds of the invention may have cytostatic activity and the sequence  
CC of the 161P2F10B protein, and the gene which encodes it, may be useful  
CC for gene therapy or the development of a vaccine. The composition and  
CC methods of the invention are useful in diagnosing, preventing and  
CC treating cancer. The present sequence is the amino acid sequence of the  
CC human 161P2F10B variant 5 protein of the invention.  
XX  
SQ Sequence 875 AA;  
  
Query Match 100.0%; Score 4804; DB 7; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVINSGLGLGLRKLKLEKQSGCRKXCFD 60  
DB 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVINSGLGLGLRKLKLEKQSGCRKXCFD 60  
  
QY 61 ASFRGLENCRCVACDRCGCCWDFDTCVETSTRIMCNKFRCGETRIEASLCSDDCL 120  
DB 61 ASFRGLENCRCVACDRCGCCWDFDTCVETSTRIMCNKFRCGETRIEASLCSDDCL 120  
  
QY 121 OKKDCADYKSCVCGRTSLENCNDTAQSQCPGDFLPPVILFSDMGFRAEYLYTDTL 180  
DB 121 OKKDCADYKSCVCGRTSLENCNDTAQSQCPGDFLPPVILFSDMGFRAEYLYTDTL 180  
  
QY 181 MPNINKLTCGHSKYNRAMPYPTKTPNHYTIVTGLYPESHGIIIDNNMYDVNLNKNFSL 240  
DB 181 MPNINKLTCGHSKYNRAMPYPTKTPNHYTIVTGLYPESHGIIIDNNMYDVNLNKNFSL 240  
  
QY 241 SKQNNPAMWHGQPMWLTANYQGLKAATYFWPGSEVAINGSPSYMPYNGSVFPEERIS 300  
DB 241 SKQNNPAMWHGQPMWLTANYQGLKAATYFWPGSEVAINGSPSYMPYNGSVFPEERIS 300  
  
QY 301 TLLKWLDPKAERPRFTYTYFEEPPSSGHAGGPVSARVICALQVVDHAFGLMEGLKQRN 360  
DB 301 TLLKWLDPKAERPRFTYTYFEEPPSSGHAGGPVSARVICALQVVDHAFGLMEGLKQRN 360  
  
QY 361 LHNVCNIIILLADHGMDOCTYCNMEYMTDYFPRINFEYVYEGAPRIRAHNIDHDFSPNS 420  
DB 361 LHNVCNIIILLADHGMDOCTYCNMEYMTDYFPRINFEYVYEGAPRIRAHNIDHDFSPNS 420  
  
QY 421 EEIVRLNSCRKPPQHKFPLTDPDLKRLHYAKNVRIDKVLHFDQDLAVRSKSNTCGG 480  
DB 421 EEIVRLNSCRKPPQHKFPLTDPDLKRLHYAKNVRIDKVLHFDQDLAVRSKSNTCGG 480  
  
QY 481 GNEGYNNEFRSMEAIPLAGPSPKETEVEPENIEVYNLMCDLRIQAPNNGTHGSLN 540  
DB 481 GNEGYNNEFRSMEAIPLAGPSPKETEVEPENIEVYNLMCDLRIQAPNNGTHGSLN 540  
  
QY 541 HLLKVPFYPSPSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTOLEQVNMNLNLTQEEI 600  
DB 541 HLLKVPFYPSPSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTOLEQVNMNLNLTQEEI 600

QY 601 TATVKVNLPEGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPLGDTSLPPT 660  
DB 601 TATVKVNLPEGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPLGDTSLPPT 660  
  
QY 661 VPCLRADVRVPSPESQKCSFYLDKKNITHGFLYPPASNRTSDSQYDALITSNLVEMYSE 720  
DB 661 VPCLRADVRVPSPESQKCSFYLDKKNITHGFLYPPASNRTSDSQYDALITSNLVEMYSE 720  
  
QY 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDVNDGHFDAPDEITKHLANTDVPPIPTHY 780  
DB 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDVNDGHFDAPDEITKHLANTDVPPIPTHY 780  
  
QY 781 FVVLTSCKNKSHTPCPGWLVDLPFIIPHRPNNVSCPEGKPEALMWVERFTAHIARVR 840  
DB 781 FVVLTSCKNKSHTPCPGWLVDLPFIIPHRPNNVSCPEGKPEALMWVERFTAHIARVR 840  
  
QY 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFTTI 875  
DB 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFTTI 875  
  
RESULT 7  
ADE65983  
ID ADE65983 standard; protein; 875 AA.  
XX  
AC ADE65983;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human 161P2F10B protein variant 1 #5.  
XX  
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003040340-A2.  
XX  
PD 15-MAY-2003.  
XX  
PF 07-NOV-2002; 2002WO-US036002.  
XX  
PR 07-NOV-2001; 2001US-00005480.  
PR 31-JAN-2002; 2002US-00062109.  
XX  
PA (AGEN-) AGENSYS INC.  
XX  
PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KM;  
PI Morrison RK, Challita-Eid PM;  
XX  
DR WPI: 2003-441560/41.  
XX  
PT A composition for diagnosing, preventing and treating cancer (e.g.  
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
PT and polypeptides.  
XX  
PS Example 5; SEQ ID NO 95; 135pp; English.  
XX  
CC This invention relates to a novel composition which comprises a substance  
CC that modulates the status of a novel protein (161P2F10B) and its variants  
CC having a sequence of 875 amino acids provided in the specification. The  
CC protein of the invention is over-expressed in certain cancers. The  
CC compounds of the invention may have cytostatic activity and the sequence  
CC of the 161P2F10B protein, and the gene which encodes it, may be useful  
CC for gene therapy or the development of a vaccine. The composition and  
CC methods of the invention are useful in diagnosing, preventing and  
CC treating cancer. The present sequence is the amino acid sequence of the  
CC human 161P2F10B variant 1 protein of the invention.  
XX  
SQ Sequence 875 AA;  
  
Query Match 100.0%; Score 4804; DB 7; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1	MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
Db 1	MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
QY 61	ASFRGLENCRCVACDRGDCWDFEDTCVSTRIMWCKNFKCGTRLEASLCSDDCL 120
Db 61	ASFRGLENCRCVACDRGDCWDFEDTCVSTRIMWCKNFKCGTRLEASLCSDDCL 120
QY 121	QKDCADYKSVCOGETSWLENCDTAQOQCPGFDLPVILFSDGFRAEYLTYWDTL 180
Db 121	QKDCADYKSVCOGETSWLENCDTAQOQCPGFDLPVILFSDGFRAEYLTYWDTL 180
QY 181	MPNINKLTCGISHKYNRMYPKTTPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
Db 181	MPNINKLTCGISHKYNRMYPKTTPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
QY 241	SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPSYMPYNGSVPEERIS 300
Db 241	SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPSYMPYNGSVPEERIS 300
QY 301	TLLKWLDPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQ 360
Db 301	TLLKWLDPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQ 360
QY 361	LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMGAPRAIRAHNIHDFSPNS 420
Db 361	LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMGAPRAIRAHNIHDFSPNS 420
QY 421	EETVRLNSCRKPDQHPKPYLTPDKRLHYAKNVRIDKVLHFDVQOVLAVRSKNTCGG 480
Db 421	EETVRLNSCRKPDQHPKPYLTPDKRLHYAKNVRIDKVLHFDVQOVLAVRSKNTCGG 480
QY 481	GHGYNNRPSRMEAFIPLAGPFGKTEVEPEENTEVNLMCDLLRIQAPNNGTHGSLN 540
Db 481	GHGYNNRPSRMEAFIPLAGPFGKTEVEPEENTEVNLMCDLLRIQAPNNGTHGSLN 540
QY 541	HLLKVPFYPSPHABEYSKSVCGFANPLPTESLDQCPHLQNSTOLEQVNMNLTOBEI 600
Db 541	HLLKVPFYPSPHABEYSKSVCGFANPLPTESLDQCPHLQNSTOLEQVNMNLTOBEI 600
QY 601	TATVKVNLFPGRPRVLQKNVDHCLLVHREYVSGFGKAMRMPWSSYTPQLGDTSPLPPT 660
Db 601	TATVKVNLFPGRPRVLQKNVDHCLLVHREYVSGFGKAMRMPWSSYTPQLGDTSPLPPT 660
QY 661	VPDCLRADYRVPPSESKCSFYLDKNIITHGFLYPPASNRSDSOYDALITSNLPMYEE 720
Db 661	VPDCLRADYRVPPSESKCSFYLDKNIITHGFLYPPASNRSDSOYDALITSNLPMYEE 720
QY 721	FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDYDGHFDAPDEITKHLANTDVPITHY 780
Db 721	FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDYDGHFDAPDEITKHLANTDVPITHY 780
QY 781	FVVLTSCKNKSHTPENCPOWLVLPPIIHRPTNVESCEGKPEALWVEERTTAHIAVR 840
Db 781	FVVLTSCKNKSHTPENCPOWLVLPPIIHRPTNVESCEGKPEALWVEERTTAHIAVR 840
QY 841	DVELLITGLDFYQKQVQPVSEILQLKTYLPTFETTI 875
Db 841	DVELLITGLDFYQKQVQPVSEILQLKTYLPTFETTI 875

RESULT 8  
AD265982  
ID ADE65982 standard; protein; 875 AA.  
XX  
AC ADE65982;  
XX  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human 161P2F10B protein variant 6 #2.  
XX

KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.	
XX Homo sapiens.	
FN WO2003040340-A2.	
XX 15-MAY-2003.	
PD 07-NOV-2002; 2002WO-US036002.	
XX 07-NOV-2001; 2001US-00005480.	
PF 31-JAN-2002; 2002US-00062109.	
XX (AGEN-) AGENSYS INC.	
PA Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;	
XX Morrison R, Challita-Eid PM;	
PI WPI; 2003-441560/41.	
DR A composition for diagnosing, preventing and treating cancer (e.g.	
XX prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides	
PT and polypeptides.	
PT Example 5; SEQ ID NO 94; 135pp; English.	
XX This invention relates to a novel composition which comprises a substance	
CC that modulates the status of a novel protein (161P2F10B) and its variants	
CC having a sequence of 875 amino acids provided in the specification. The	
CC protein of the invention is over-expressed in certain cancers. The	
CC compounds of the invention may have cytostatic activity and the sequence	
CC of the 161P2F10B protein, and the gene which encodes it, may be useful	
CC for gene therapy or the development of a vaccine. The composition and	
CC methods of the invention are useful in diagnosing, preventing and	
CC treating cancer. The present sequence is the amino acid sequence of the	
CC human 161P2F10B variant 6 protein of the invention.	
XX Sequence 875 AA;	
Query Match 100.0%; Score 4804; DB 7; Length 875;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1	MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
Db 1	MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
QY 61	ASFRGLENCRCVACDRGDCWDFEDTCVSTRIMWCKNFKCGTRLEASLCSDDCL 120
Db 61	ASFRGLENCRCVACDRGDCWDFEDTCVSTRIMWCKNFKCGTRLEASLCSDDCL 120
QY 121	QKDCADYKSVCOGETSWLENCDTAQOQCPGFDLPVILFSDGFRAEYLTYWDTL 180
Db 121	QKDCADYKSVCOGETSWLENCDTAQOQCPGFDLPVILFSDGFRAEYLTYWDTL 180
QY 181	MPNINKLTCGISHKYNRMYPKTTPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
Db 181	MPNINKLTCGISHKYNRMYPKTTPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
QY 241	SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPSYMPYNGSVPEERIS 300
Db 241	SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPSYMPYNGSVPEERIS 300
QY 301	TLLKWLDPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQ 360
Db 301	TLLKWLDPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQ 360
QY 361	LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMGAPRAIRAHNIHDFSPNS 420
Db 361	LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMGAPRAIRAHNIHDFSPNS 420
QY 421	EETVRLNSCRKPDQHPKPYLTPDKRLHYAKNVRIDKVLHFDVQOVLAVRSKNTCGG 480
Db 421	EETVRLNSCRKPDQHPKPYLTPDKRLHYAKNVRIDKVLHFDVQOVLAVRSKNTCGG 480
QY 481	GHGYNNRPSRMEAFIPLAGPFGKTEVEPEENTEVNLMCDLLRIQAPNNGTHGSLN 540
Db 481	GHGYNNRPSRMEAFIPLAGPFGKTEVEPEENTEVNLMCDLLRIQAPNNGTHGSLN 540
QY 541	HLLKVPFYPSPHABEYSKSVCGFANPLPTESLDQCPHLQNSTOLEQVNMNLTOBEI 600
Db 541	HLLKVPFYPSPHABEYSKSVCGFANPLPTESLDQCPHLQNSTOLEQVNMNLTOBEI 600
QY 601	TATVKVNLFPGRPRVLQKNVDHCLLVHREYVSGFGKAMRMPWSSYTPQLGDTSPLPPT 660
Db 601	TATVKVNLFPGRPRVLQKNVDHCLLVHREYVSGFGKAMRMPWSSYTPQLGDTSPLPPT 660
QY 661	VPDCLRADYRVPPSESKCSFYLDKNIITHGFLYPPASNRSDSOYDALITSNLPMYEE 720
Db 661	VPDCLRADYRVPPSESKCSFYLDKNIITHGFLYPPASNRSDSOYDALITSNLPMYEE 720
QY 721	FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDYDGHFDAPDEITKHLANTDVPITHY 780
Db 721	FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDYDGHFDAPDEITKHLANTDVPITHY 780
QY 781	FVVLTSCKNKSHTPENCPOWLVLPPIIHRPTNVESCEGKPEALWVEERTTAHIAVR 840
Db 781	FVVLTSCKNKSHTPENCPOWLVLPPIIHRPTNVESCEGKPEALWVEERTTAHIAVR 840
QY 841	DVELLITGLDFYQKQVQPVSEILQLKTYLPTFETTI 875
Db 841	DVELLITGLDFYQKQVQPVSEILQLKTYLPTFETTI 875

Db 421 BEIVRNLSCKRPQHPKPYLTDPDLKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCGG 480  
 QY 481 GNGYNNFRSMEAI FLAHGSPFKETVEPEFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 Db 481 GNGYNNFRSMEAI FLAHGSPFKETVEPEFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQLNSTOLEQVQMLNLTQBEI 600  
 Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQLNSTOLEQVQMLNLTQBEI 600  
 QY 601 TATVKVNLPEGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 Db 601 TATVKVNLPEGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 QY 661 VPCLRADVRVPSESKCSFYLDKNITHGFLYPASNRTSDSQDALITSNLVPMYEE 720  
 Db 661 VPCLRADVRVPSESKCSFYLDKNITHGFLYPASNRTSDSQDALITSNLVPMYEE 720  
 QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780  
 Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780  
 QY 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIARVR 840  
 Db 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIARVR 840  
 QY 841 DVELLTGLDFYQDKVPVSEILQKTYLPTFTTI 875  
 Db 841 DVELLTGLDFYQDKVPVSEILQKTYLPTFTTI 875

RESULT 9  
 ADE65909 standard; protein; 875 AA.  
 AC ADE65909;  
 DT 29-JAN-2004 (first entry)  
 XX Human 161P2F10B protein variant 1 #3.  
 DE 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.  
 XX Homo sapiens.  
 OS WO2003040340-A2.  
 XX 15-MAY-2003.  
 PD 07-NOV-2002; 2002WO-US036002.  
 PF 07-NOV-2001; 2001US-00005480.  
 XX 31-JAN-2002; 2002US-00062109.  
 PR (AGEN-) AGENSYS INC.  
 PA Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KM;  
 PI Morrison RK, Chailita-Eid PM;  
 XX WPI; 2003-441560/41.  
 DR A composition for diagnosing, preventing and treating cancer (e.g.  
 PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
 PT and polypeptides.  
 XX Example 2; SEQ ID NO 21; 135pp; English.  
 PS This invention relates to a novel composition which comprises a substance  
 CC that modulates the status of a novel protein (161P2F10B) and its variants  
 CC having a sequence of 875 amino acids provided in the specification. The  
 CC protein of the invention is over-expressed in certain cancers. The  
 CC compounds of the invention may have cytostatic activity and the sequence  
 CC of the 161P2F10B protein, and the gene which encodes it, may be useful

CC for gene therapy or the development of a vaccine. The composition and  
 CC methods of the invention are useful in diagnosing, preventing and  
 CC treating cancer. The present sequence is the amino acid sequence of the  
 CC human 161P2F10B variant 1 protein of the invention.  
 XX  
 SQ Sequence 875 AA;  
 Query Match 100.0%; Score 4804; DB 7; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MESTLTILATEQPVKNTLKKYKIACIVLLALLVIMSLGIGLGLKRLKQSGCRKKCFD 60  
 Db 1 MESTLTILATEQPVKNTLKKYKIACIVLLALLVIMSLGIGLGLKRLKQSGCRKKCFD 60  
 QY 61 ASFRGLENCRCVACDRGDCWDFEDTCVESTRIWNCNKRFCGETRLEASLCSCSDDC 120  
 Db 61 ASFRGLENCRCVACDRGDCWDFEDTCVESTRIWNCNKRFCGETRLEASLCSCSDDC 120  
 QY 121 OKDCCADYKSVCOGETSMLNENCDAQOSQCEGDLPPVILFSDMGFAEVLITWDTL 180  
 Db 121 OKDCCADYKSVCOGETSMLNENCDAQOSQCEGDLPPVILFSDMGFAEVLITWDTL 180  
 QY 181 MPNINKLKTGCIHSKYMRAVYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240  
 Db 181 MPNINKLKTGCIHSKYMRAVYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240  
 QY 241 SKQNNPAWHGQPMMLTANYQGLKAATYFWPGEVAINGSPSIIYMPYNGSVPPERIS 300  
 Db 241 SKQNNPAWHGQPMMLTANYQGLKAATYFWPGEVAINGSPSIIYMPYNGSVPPERIS 300  
 QY 301 TLLKWLDPKAEPRFYTMYFEEPDSSGHAGGSARVVKALQVVDHAFGLMEGLKQEN 360  
 Db 301 TLLKWLDPKAEPRFYTMYFEEPDSSGHAGGSARVVKALQVVDHAFGLMEGLKQEN 360  
 QY 361 LHCNVIILLADHGMDOQTYCNKMEYMTDYPFRINPFYMEGAPAPRIRAINIIPHDFFS 420  
 Db 361 LHCNVIILLADHGMDOQTYCNKMEYMTDYPFRINPFYMEGAPAPRIRAINIIPHDFFS 420  
 QY 421 BEIVRNLSCKRPQHPKPYLTDPDLKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCGG 480  
 Db 421 BEIVRNLSCKRPQHPKPYLTDPDLKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNTCGG 480  
 QY 481 GNGYNNFRSMEAI FLAHGSPFKETVEPEFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 Db 481 GNGYNNFRSMEAI FLAHGSPFKETVEPEFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQLNSTOLEQVQMLNLTQBEI 600  
 Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQLNSTOLEQVQMLNLTQBEI 600  
 QY 601 TATVKVNLPEGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 Db 601 TATVKVNLPEGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 QY 661 VPCLRADVRVPSESKCSFYLDKNITHGFLYPASNRTSDSQDALITSNLVPMYEE 720  
 Db 661 VPCLRADVRVPSESKCSFYLDKNITHGFLYPASNRTSDSQDALITSNLVPMYEE 720  
 QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780  
 Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780  
 QY 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIARVR 840  
 Db 781 FVVLTSCKNKSHTPENCPCGMLDVLPIIIPHRPTNVESCEPGKPEALWVEERFTHIARVR 840  
 QY 841 DVELLTGLDFYQDKVPVSEILQKTYLPTFTTI 875  
 Db 841 DVELLTGLDFYQDKVPVSEILQKTYLPTFTTI 875

RESULT 10



XX Disclosure; Fig 13A; 135pp; English.

XX This invention relates to a novel composition which comprises a substance that modulates the status of a novel protein (161P2F10B) and its variants having a sequence of 875 amino acids provided in the specification. The protein of the invention is over-expressed in certain cancers. The compounds of the invention may have cytostatic activity and the sequence of the 161P2F10B protein, and the gene which encodes it, may be useful for gene therapy or the development of a vaccine. The composition and methods of the invention are useful in diagnosing, preventing and treating cancer. The present sequence is the amino acid sequence of the human 161P2F10B variant 1 protein of the invention.

XX Sequence 875 AA;

XX Query Match 100.0%; Score 4804; DB 7; Length 875;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESTLTATEQPVKNTLKKYKIACIALLVIMSLGLGLGRKLEKQSCSKKCFD 60

Db 1 MESTLTATEQPVKNTLKKYKIACIALLVIMSLGLGLGRKLEKQSCSKKCFD 60

Qy 61 ASFRGLENCRCVACKDRGDCWDVEDTCTVESTRIMWCKNFKCGTRLEASI-CSSDDCL 120

Db 61 ASFRGLENCRCVACKDRGDCWDVEDTCTVESTRIMWCKNFKCGTRLEASLCSSDDCL 120

Qy 121 QKDCCADYKSCQGETSWLENCDTAQQSCPEGFDLPVILFMSMDGFRAEYLTYDNL 180

Db 121 QKDCCADYKSCQGETSWLENCDTAQQSCPEGFDLPVILFMSMDGFRAEYLTYDNL 180

Qy 181 MNINKLTCGTHSKYRAMPYTKFPNHYTIVTGLYPESHGIIIDNNYDNLNKNFSL 240

Db 181 MNINKLTCGTHSKYRAMPYTKFPNHYTIVTGLYPESHGIIIDNNYDNLNKNFSL 240

Qy 241 SKEQNPANWHGQPMWLTAMYQGLKAATYFPGSEVAINGSPFSIYMPYNGSVPEERIS 300

Db 241 SKEQNPANWHGQPMWLTAMYQGLKAATYFPGSEVAINGSPFSIYMPYNGSVPEERIS 300

Qy 301 TLKWLKDLPAERPRYTYFPEPSSHAGCPVSARVIALQVDFHAFGLMEGLKQRN 360

Db 301 TLKWLKDLPAERPRYTYFPEPSSHAGCPVSARVIALQVDFHAFGLMEGLKQRN 360

Qy 361 LHCNVNILLADHGMDOYTCNKMEYMTDYFRINFEYVYEGPAPRAIRAHNIPHDFFSNS 420

Db 361 LHCNVNILLADHGMDOYTCNKMEYMTDYFRINFEYVYEGPAPRAIRAHNIPHDFFSNS 420

Qy 421 EEIVRLNLSCKRQDPQHPKPYLTPDLKRLHYAKNRIDKVLHFDQOQLAVRSKNTNCGG 480

Db 421 EEIVRLNLSCKRQDPQHPKPYLTPDLKRLHYAKNRIDKVLHFDQOQLAVRSKNTNCGG 480

Qy 481 GNHGYNNEFRSMEAIPLANGPSEKTEVEPEPENTVNLNCDLRIOPAPNNGTHGSLN 540

Db 481 GNHGYNNEFRSMEAIPLANGPSEKTEVEPEPENTVNLNCDLRIOPAPNNGTHGSLN 540

Qy 541 HLLKVPFYEPSSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTOLEQVQNLNLTQEEI 600

Db 541 HLLKVPFYEPSSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTOLEQVQNLNLTQEEI 600

Qy 601 TATVKNLDPGPRVLQKNVDHCLLYHREYVSGFKAMRMPWSSYTYPQLGDTSPILPT 660

Db 601 TATVKNLDPGPRVLQKNVDHCLLYHREYVSGFKAMRMPWSSYTYPQLGDTSPILPT 660

Qy 661 VPDCLRADVRVPSSQKCSFYLDKNIITHGLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db 661 VPDCLRADVRVPSSQKCSFYLDKNIITHGLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy 721 FRKMDYFHSVLLIKHATERGNVNVSGPIFDYNDGHFDAPDEITKGLANTDVPITPHY 780

Db 721 FRKMDYFHSVLLIKHATERGNVNVSGPIFDYNDGHFDAPDEITKGLANTDVPITPHY 780

Qy 781 FVVLTSCKNKSHTPENCPCGMLDVLFFIIPHRPTNVESCPGKPEALWVEERFTHAIAVR 840

Db 781 FVVLTSCKNKSHTPENCPCGMLDVLFFIIPHRPTNVESCPGKPEALWVEERFTHAIAVR 840

Qy 841 DVELLTGLDFYQDKVQVSVSEILOLKYLYPTFETI 875

Db 841 DVELLTGLDFYQDKVQVSVSEILOLKYLYPTFETI 875

RESULT 12

AD65910

ID AD65910 standard; protein; 875 AA.

XX AC AD65910;

XX 29-JAN-2004 (first entry)

XX Human 161P2F10B protein K122R variant.

XX 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.

XX Homo sapiens.

XX WO2003040340-A2.

XX 15-MAY-2003.

XX 07-NOV-2002; 2002WO-US036002.

XX 07-NOV-2001; 2001US-00005480.

XX 31-JAN-2002; 2002US-00062109.

XX (AGEN-) AGENSYS INC.

XX Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KM;

XX Morrison RK, Challita-Eid PM;

XX WPI; 2003-441560/41.

XX A composition for diagnosing, preventing and treating cancer (e.g. prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides and polypeptides.

XX Example 2; SEQ ID NO 22; 135pp; English.

XX This invention relates to a novel composition which comprises a substance that modulates the status of a novel protein (161P2F10B) and its variants having a sequence of 875 amino acids provided in the specification. The protein of the invention is over-expressed in certain cancers. The compounds of the invention may have cytostatic activity and the sequence of the 161P2F10B protein, and the gene which encodes it, may be useful for gene therapy or the development of a vaccine. The composition and methods of the invention are useful in diagnosing, preventing and treating cancer. The present sequence is the amino acid sequence of the human 161P2F10B K122R variant protein which was used for homology purposes and is related to the invention.

XX Sequence 875 AA;

XX Query Match 99.9%; Score 4801; DB 7; Length 875;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 874; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESTLTATEQPVKNTLKKYKIACIALLVIMSLGLGLGRKLEKQSCSKKCFD 60

Db 1 MESTLTATEQPVKNTLKKYKIACIALLVIMSLGLGLGRKLEKQSCSKKCFD 60

Qy 61 ASFRGLENCRCVACKDRGDCWDVEDTCTVESTRIMWCKNFKCGTRLEASI-CSSDDCL 120

Db 61 ASFRGLENCRCVACKDRGDCWDVEDTCTVESTRIMWCKNFKCGTRLEASLCSSDDCL 120

Qy 121 QKDCCADYKSCQGETSWLENCDTAQQSCPEGFDLPVILFMSMDGFRAEYLTYDNL 180

Db 121 QKDCCADYKSCQGETSWLENCDTAQQSCPEGFDLPVILFMSMDGFRAEYLTYDNL 180

QY 181 MPNINKLKTGIIHSKYRMYPTKTPFNHYTIVTGLYPESHGIIIDNNMVDVNLKNFSL 240  
 Db 181 MPNINKLKTGIIHSKYRMYPTKTPFNHYTIVTGLYPESHGIIIDNNMVDVNLKNFSL 240  
 QY 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300  
 Db 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300  
 QY 301 TLLKWLDPKAEPRFRFTMYFEPPDSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
 Db 301 TLLKWLDPKAEPRFRFTMYFEPPDSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360  
 QY 361 LHCNVNIIILLADHGMDQTYCNKMEYMTDYPFRINFFMYEGPAPRIAHNIHDPFSFNS 420  
 Db 361 LHCNVNIIILLADHGMDQTYCNKMEYMTDYPFRINFFMYEGPAPRIAHNIHDPFSFNS 420  
 QY 421 BEIVRNLSCKPDQHFPPYLTDPDLKRLHYAKNVRIDKVLHFDVQDLAVRSKSNNTCCG 480  
 Db 421 BEIVRNLSCKPDQHFPPYLTDPDLKRLHYAKNVRIDKVLHFDVQDLAVRSKSNNTCCG 480  
 QY 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 Db 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
 QY 541 HLLKVPFYPESHAEVSKSVCGFANPLPTESLDCTCPHLQNSQLEQVNMNLTOBEI 600  
 Db 541 HLLKVPFYPESHAEVSKSVCGFANPLPTESLDCTCPHLQNSQLEQVNMNLTOBEI 600  
 QY 601 TATVKVNLPGFRVQLQNDHCLLYHREYVSGFKAMRPMWSSYTVPLQDTSPLPT 660  
 Db 601 TATVKVNLPGFRVQLQNDHCLLYHREYVSGFKAMRPMWSSYTVPLQDTSPLPT 660  
 QY 661 VPDCLRADVRVPSESOKCSFYADKNITHGFLYPPASNRTSDSYDALITSNLVPMYEE 720  
 Db 661 VPDCLRADVRVPSESOKCSFYADKNITHGFLYPPASNRTSDSYDALITSNLVPMYEE 720  
 QY 721 FRKMDYFHSVLLIKHATERGVNVVSGPIFDYNDGHPDAPDEITKHLANTDVPIPTHY 780  
 Db 721 FRKMDYFHSVLLIKHATERGVNVVSGPIFDYNDGHPDAPDEITKHLANTDVPIPTHY 780  
 QY 781 FVLTSCNKSHTPENCGLDVLFFIIPHRPTNVSCPEGKPEALWBERFTHIARVR 840  
 Db 781 FVLTSCNKSHTPENCGLDVLFFIIPHRPTNVSCPEGKPEALWBERFTHIARVR 840  
 QY 841 DVELLTGLDFQDKVQVPSILQIKTYLPTFFETI 875  
 Db 841 DVELLTGLDFQDKVQVPSILQIKTYLPTFFETI 875

## RESULT 13

ADE65905

ID ADE65905 standard; protein; 875 AA.

XX AC ADE65905;

XX AC ADE65905;

XX DT 29-JAN-2004 (first entry)

XX DE Human 161P2F10B protein variant 2 #2.

XX KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.

XX OS Homo sapiens.

XX FN WO2003040340-A2.

XX PD 15-MAY-2003.

XX XX 07-NOV-2002; 2002WO-US036002.

XX XX 07-NOV-2001; 2001US-00005480.

XX PR 31-JAN-2002; 2002US-00062109.

XX XX

## (AGEN-) AGENSYS INC.

Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KM;

Morrison RK, Challita-Eid PN;

WPI; 2003-441560/41.

DR N-PSDB; ADE65893.

XX

A composition for diagnosing, preventing and treating cancer (e.g.

prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides

and polypeptides.

XX

Example 1; SEQ ID NO 17; 135pp; English.

PS

This invention relates to a novel composition which comprises a substance

that modulates the status of a novel protein (161P2F10B) and its variants

having a sequence of 875 amino acids provided in the specification. The

protein of the invention is over-expressed in certain cancers. The

compounds of the invention may have cytostatic activity and the sequence

of the 161P2F10B protein, and the gene which encodes it, may be useful

for gene therapy or the development of a vaccine. The composition and

methods of the invention are useful in diagnosing, preventing and

treating cancer. The present sequence is the amino acid sequence of the

human 161P2F10B variant 2 protein of the invention.

XX

SQ Sequence 875 AA;

Query Match

Best Local Similarity 99.9%; Score 4801; DB 7; Length 875;

Matches 874; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATQPKVQKNTLKYKIACIVLLALLVMSLGLGLGLKLEKQSKCKKCFD 60

Db 1 MESTLTATQPKVQKNTLKYKIACIVLLALLVMSLGLGLGLKLEKQSKCKKCFD 60

QY 61 ASFRGLENCDCVACKDRCDCDCCDFEDTCTVETRIWCKNFKRCGETELASLCSDDCL 120

Db 61 ASFRGLENCDCVACKDRCDCDCCDFEDTCTVETRIWCKNFKRCGETELASLCSDDCL 120

QY 121 QKDCADYKSVCGGETSWLENCDTAQSCQCEGFDLPVILFSDMGFRAEYLYTMDTL 180

Db 121 QKDCADYKSVCGGETSWLENCDTAQSCQCEGFDLPVILFSDMGFRAEYLYTMDTL 180

QY 181 MPNINKLKTGIIHSKYRMYPTKTPFNHYTIVTGLYPESHGIIIDNNMVDVNLKNFSL 240

Db 181 MPNINKLKTGIIHSKYRMYPTKTPFNHYTIVTGLYPESHGIIIDNNMVDVNLKNFSL 240

QY 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300

Db 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300

QY 301 TLLKWLDPKAEPRFRFTMYFEPPDSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360

Db 301 TLLKWLDPKAEPRFRFTMYFEPPDSGHAGGVPVSARVIALQVVDHAFGLMEGLKQRN 360

QY 361 LHCNVNIIILLADHGMDQTYCNKMEYMTDYPFRINFFMYEGPAPRIAHNIHDPFSFNS 420

Db 361 LHCNVNIIILLADHGMDQTYCNKMEYMTDYPFRINFFMYEGPAPRIAHNIHDPFSFNS 420

QY 421 BEIVRNLSCKPDQHFPPYLTDPDLKRLHYAKNVRIDKVLHFDVQDLAVRSKSNNTCCG 480

Db 421 BEIVRNLSCKPDQHFPPYLTDPDLKRLHYAKNVRIDKVLHFDVQDLAVRSKSNNTCCG 480

QY 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540

Db 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540

QY 541 HLLKVPFYPESHAEVSKSVCGFANPLPTESLDCTCPHLQNSQLEQVNMNLTOBEI 600

Db 541 HLLKVPFYPESHAEVSKSVCGFANPLPTESLDCTCPHLQNSQLEQVNMNLTOBEI 600

QY 601 TATVKVNLPGFRVQLQNDHCLLYHREYVSGFKAMRPMWSSYTVPLQDTSPLPT 660

Db 601 TATVKVNLPGFRVQLQNDHCLLYHREYVSGFKAMRPMWSSYTVPLQDTSPLPT 660



Db 601 TATVKNLPGFRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 Qy 661 VPDCLRADVRVPSESKCSFYLDKNIITHGFLYPPASNRTSDSYDALITSLNVPWYEE 720  
 Db 661 VPDCLRADVRVPSESKCSFYLDKNIITHGFLYPPASNRTSDSYDALITSLNVPWYEE 720  
 Qy 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDYGHFDAPDEITKHLANTDVPITPHY 780  
 Db 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDYGHFDAPDEITKHLANTDVPITPHY 780  
 Qy 781 FVVLTSCKNKSHTPENCPCWGLDVLPIIPIHRTNVSCPEGKPEALWVEERFTHAIRVR 840  
 Db 781 FVVLTSCKNKSHTPENCPCWGLDVLPIIPIHRTNVSCPEGKPEALWVEERFTHAIRVR 840  
 Qy 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875  
 Db 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875

## RESULT 14

AD65893  
 ID ADE65893 standard; protein; 875 AA.

AC ADE65893;

DT 29-JAN-2004 (first entry)

DE Human 161P2F10B protein variant 2.

XX 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.

OS Homo sapiens.

PN WC2003040340-A2.

PD 15-MAY-2003.

PF 07-NOV-2002; 2002WO-US036002.

PR 07-NOV-2003; 2001US-00005480.

PR 31-JAN-2002; 2002US-00062109.

XX (AGEN-) AGENSYS INC.

PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;  
 PI Morrison RK, Challita-Bid PM;

DR WPI; 2003-441560/41.

DR N-PSDB; AD65892.

XX A composition for diagnosing, preventing and treating cancer (e.g.  
 PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
 PT and polypeptides.

PS Claim 1; SEQ ID NO 5; 135pp; English.

XX This invention relates to a novel composition which comprises a substance  
 CC that modulates the status of a novel protein (161P2F10B) and its variants  
 CC having a sequence of 875 amino acids provided in the specification. The  
 CC protein of the invention is over-expressed in certain cancers. The  
 CC compounds of the invention may have cytostatic activity and the sequence  
 CC of the 161P2F10B protein, and the gene which encodes it, may be useful  
 CC for gene therapy or the development of a vaccine. The composition and  
 CC methods of the invention are useful in diagnosing, preventing and  
 CC treating cancer. The present sequence is the amino acid sequence of the  
 CC human 161P2F10B variant 2 protein of the invention.

XX Sequence 875 AA;

Query Match

Best Local Similarity 99.9%; Score 4801; DB 7; Length 875;

Matches 874; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESTLTLTATEQPVKQNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60  
 Db 1 MESTLTLTATEQPVKQNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60  
 Qy 61 ASFRGLENCRCDAKCRGDCWDPEDTCVESTETIWNCKNRCGETRLEASLCSDDCL 120  
 Db 61 ASFRGLENCRCDAKCRGDCWDPEDTCVESTETIWNCKNRCGETRLEASLCSDDCL 120  
 Qy 121 QKQCCADYKSVCCGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEVLVYTDITL 180  
 Db 121 QKQCCADYKSVCCGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEVLVYTDITL 180  
 Qy 181 MPNINKLKTGIIHSKYRAMPYTKTFPNHYTIVTGLYPESHGIIIDNNMYDNLNKPFLS 240  
 Db 181 MPNINKLKTGIIHSKYRAMPYTKTFPNHYTIVTGLYPESHGIIIDNNMYDNLNKPFLS 240  
 Qy 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPFPERIS 300  
 Db 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPFPERIS 300  
 Qy 301 TLLKWLDPKAEPRFTMYFEEPDSSGHAGGVPVSARVIKALQVDFHAFQMLMEGLKQRN 360  
 Db 301 TLLKWLDPKAEPRFTMYFEEPDSSGHAGGVPVSARVIKALQVDFHAFQMLMEGLKQRN 360  
 Qy 361 LHCNVIILLADHGMDOYCNKMEYMTDYPFRINFFYMEGAPAPRIRAHNI PHDFSFS 420  
 Db 361 LHCNVIILLADHGMDOYCNKMEYMTDYPFRINFFYMEGAPAPRIRAHNI PHDFSFS 420  
 Qy 421 BEIVRNLSCKRPQHFVLTDPDLPKRLHYAKNVRIDKVHLFVDOQWLAVRSKSNINCG 480  
 Db 421 BEIVRNLSCKRPQHFVLTDPDLPKRLHYAKNVRIDKVHLFVDOQWLAVRSKSNINCG 480  
 Qy 481 GNEGYNNPESMEAI FLAAGPSFKETEPENIEVNLMDLRIQAPNNGTHGSLN 540  
 Db 481 GNEGYNNPESMEAI FLAAGPSFKETEPENIEVNLMDLRIQAPNNGTHGSLN 540  
 Qy 541 HLLKVPFYPFSAEVSKEFVCGFANPLPTESIDCFPHLQNSTOLEQVNMNLTOEII 600  
 Db 541 HLLKVPFYPFSAEVSKEFVCGFANPLPTESIDCFPHLQNSTOLEQVNMNLTOEII 600  
 Qy 601 TATVKVNLPGFRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 Db 601 TATVKVNLPGFRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660  
 Qy 661 VPDCLRADVRVPSESKCSFYLDKNIITHGFLYPPASNRTSDSYDALITSLNVPWYEE 720  
 Db 661 VPDCLRADVRVPSESKCSFYLDKNIITHGFLYPPASNRTSDSYDALITSLNVPWYEE 720  
 Qy 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDYGHFDAPDEITKHLANTDVPITPHY 780  
 Db 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDYGHFDAPDEITKHLANTDVPITPHY 780  
 Qy 781 FVVLTSCKNKSHTPENCPCWGLDVLPIIPIHRTNVSCPEGKPEALWVEERFTHAIRVR 840  
 Db 781 FVVLTSCKNKSHTPENCPCWGLDVLPIIPIHRTNVSCPEGKPEALWVEERFTHAIRVR 840  
 Qy 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875  
 Db 841 DVELLTGLDFYQDKVQVPVSEILQLKTYLPTFTTI 875

## RESULT 15

AD65897

ID ADE65897 standard; protein; 875 AA.

XX ADE65897;

XX 29-JAN-2004 (first entry)

XX Human 161P2F10B protein variant 4.

XX 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.



OS Homo sapiens.  
XX WO2003040340-A2.  
XX 15-MAY-2003.  
XX 07-NOV-2002; 2002WO-US036002.  
XX 07-NOV-2001; 2001US-00005480.  
XX 31-JAN-2002; 2002US-00062109.  
XX (AGEN-) AGENSYS INC.  
XX Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;  
XX Morrison RK, Challita-Eid PM;  
XX WPI: 2003-441560/41.  
XX N-PSDB; ADE65896.  
XX  
XX A composition for diagnosing, preventing and treating cancer (e.g.  
XX for prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides  
XX and polypeptides.  
XX Claim 1; SEQ ID NO 9; 135pp; English.  
XX  
XX This invention relates to a novel composition which comprises a substance  
XX that modulates the status of a novel protein (161P2F10B) and its variants  
XX having a sequence of 875 amino acids provided in the specification. The  
XX protein of the invention is over-expressed in certain cancers. The  
XX compounds of the invention may have cytostatic activity and the sequence  
XX of the 161P2F10B protein, and the gene which encodes it, may be useful  
XX for gene therapy or the development of a vaccine. The composition and  
XX methods of the invention are useful in diagnosing, preventing and  
XX treating cancer. The present sequence is the amino acid sequence of the  
XX human 161P2F10B variant 4 protein of the invention.  
XX  
XX Sequence 875 AA;  
SQ  
Query Match 99.9%; Score 4798; DB 7; Length 875;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 874; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALVIMSLGGLGLGLKLEKQSCRRKCFD 60  
Db 1 MESTLTATEQPVKNTLKKYKIACIVLLALVIMSLGGLGLGLKLEKQSCRRKCFD 60  
QY 61 ASFRGLENCRCVACKDRGDCWDFFDTCVESTRIWMCNFKRCGETRLEASLCSDDCL 120  
Db 61 ASFRGLENCRCVACKDRGDCWDFFDTCVESTRIWMCNFKRCGETRLEASLCSDDCL 120  
QY 121 QKXCCADYKSVCOGETSWLENCDDTAQSQCPGEGDLPVILFSDGPRAEYLYTWDTL 180  
Db 121 QKXCCADYKSVCOGETSWLENCDDTAQSQCPGEGDLPVILFSDGPRAEYLYTWDTL 180  
QY 181 MPNINKLKTGHIHSKYNRAMPYTKTPFNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240  
Db 181 MPNINKLKTGHIHSKYNRAMPYTKTPFNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240  
QY 241 SKQNPAPWHGQPMWLTAMYQGLKAATYFPGSEVAINGSPSIYMPYNGSVPPFEERIS 300  
Db 241 SKQNPAPWHGQPMWLTAMYQGLKAATYFPGSEVAINGSPSIYMPYNGSVPPFEERIS 300  
QY 301 TLLKWLDPKAEPRFVTMYFEEDPSGSHAGGVPVSARVIKALQVVDHAFGLMEGLKQRN 360  
Db 301 TLLKWLDPKAEPRFVTMYFEEDPSGSHAGGVPVSARVIKALQVVDHAFGLMEGLKQRN 360  
QY 361 LHCNVTIILLADHGMOTYCNKMEYMTDYPFRNFFMYEGGAPRIAHNI PHDFFSFNS 420  
Db 361 LHCNVTIILLADHGMOTYCNKMEYMTDYPFRNFFMYEGGAPRIAHNI PHDFFSFNS 420  
QY 421 EETVRNLSCRKPDQHFKPYLTDPKRLHYAKNVRIDKVLHFVDQOWLA VRSKSNTNCGG 480  
Db 421 EETVRNLSCRKPDQHFKPYLTDPKRLHYAKNVRIDKVLHFVDQOWLA VRSKSNTNCGG 480

QY 481 GNGHYNNEFRSMEAIFLAHGPSFKKTEVEFPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Db 481 GNGHYNNEFRSMEAIFLAHGPSFKKTEVEFPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
QY 541 HLLKVPFYBPSHAEBVSFVCGFANPLPTESLDCEPHLQNSTOLEQVQMLNLTQBEI 600  
Db 541 HLLKVPFYBPSHAEBVSFVCGFANPLPTESLDCEPHLQNSTOLEQVQMLNLTQBEI 600  
QY 601 TATVKVNLFPGRPRVLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPLGDTSLPPT 660  
Db 601 TATVKVNLFPGRPRVLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPLGDTSLPPT 660  
QY 661 VPDCLRADYRVPSESQKCSFYLDKNITGHFLYPPASNRSDSOYDALITSNLVPMYEE 720  
Db 661 VPDCLRADYRVPSESQKCSFYLDKNITGHFLYPPASNRSDSOYDALITSNLVPMYEE 720  
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPITPHY 780  
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPITPHY 780  
QY 781 FVVLTSCKNKSHTPENCPCGWLDPVLPFIIPHRPTNVESCPGKPEALWBERFTAHIARVR 840  
Db 781 FVVLTSCKNKSHTPENCPCGWLDPVLPFIIPHRPTNVESCPGKPEALWBERFTAHIARVR 840  
QY 841 DVELLTGLDIFYQDKVQPVSEILQKTYLPTFTPI 875  
Db 841 DVELLTGLDIFYQDKVQPVSEILQKTYLPTFTPI 875

Search completed: July 6, 2004, 13:18:17  
Job time : 68 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:17:02 ; Search time 23 Seconds  
(without alignments)  
1964.032 Million cell updates/sec

Title: US-10-005-480A-743

Perfect score: 4804

Sequence: 1 MESTLTLATEQPVKNTLKK.....QPVSEILOKTYLPTFTTI 875

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pap:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pap:\*  
3: /cgn2\_6/prodata/2/iaa/5A COMB.pap:\*  
4: /cgn2\_6/prodata/2/iaa/5B COMB.pap:\*  
5: /cgn2\_6/prodata/2/iaa/5A COMB.pap:\*  
6: /cgn2\_6/prodata/2/iaa/5B COMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2498.5	52.0	873	3	US-09-187-331-6
2	2498.5	52.0	873	4	US-09-470-946-6
3	2498.5	52.0	873	4	US-09-438-908-2
4	2498.5	52.0	925	2	US-08-392-946-1
5	2498.5	52.0	925	2	US-08-504-169-1
6	2498.5	52.0	925	5	PCT-US94-14893-1
7	2494.5	51.9	873	4	US-09-438-906-4
8	2040.5	42.5	915	1	US-08-346-455B-69
9	2040.5	42.5	915	3	US-08-977-221-69
10	2040.5	42.5	915	4	US-09-483-831B-69
11	2040.5	42.5	915	5	PCT-US95-06613-69
12	1964.5	40.9	829	3	US-08-346-455B-34
13	1964.5	40.9	829	3	US-08-977-221-34
14	1964.5	40.9	829	4	US-09-483-831B-34
15	1964.5	40.9	829	5	PCT-US95-06613-34
16	1923.5	40.0	861	1	US-08-346-455B-67
17	1923.5	40.0	861	3	US-08-977-221-67
18	1923.5	40.0	861	4	US-09-483-831B-67
19	1923.5	40.0	861	5	PCT-US95-06613-67
20	1838.5	38.3	979	1	US-08-346-455B-38
21	1838.5	38.3	979	3	US-08-977-221-38
22	1838.5	38.3	979	4	US-09-483-831B-70
23	1838.5	38.3	979	5	PCT-US95-06613-38
24	1761.5	36.7	788	1	US-08-346-455B-36
25	1761.5	36.7	788	3	US-08-977-221-36
26	1761.5	36.7	788	4	US-09-483-831B-36
27	1761.5	36.7	788	5	PCT-US95-06613-36

28	641	13.3	453	4	US-09-800-729-83	Sequence 83, Appl
29	503	10.5	355	2	US-09-014-969-19	Sequence 19, Appl
30	432.5	9.0	438	3	US-09-187-331-2	Sequence 2, Appl
31	432.5	9.0	438	4	US-09-470-946-2	Sequence 2, Appl
32	227.5	4.7	136	4	US-09-621-976-3913	Sequence 3913, Ap
33	210.5	4.4	151	4	US-09-621-976-3891	Sequence 3891, Ap
34	144.5	3.0	1354	4	US-07-757-022B-48	Sequence 48, Appl
35	143	3.0	1022	4	US-07-757-022B-84	Sequence 84, Appl
36	140.5	2.9	937	4	US-07-757-022B-72	Sequence 72, Appl
37	139.5	2.9	1311	4	US-07-757-022B-42	Sequence 42, Appl
38	139	2.9	463	4	US-07-757-022B-54	Sequence 54, Appl
39	136.5	2.8	157	4	US-07-757-022B-102	Sequence 102, App
40	136.5	2.8	157	4	US-07-757-022B-114	Sequence 114, App
41	134	2.8	106	4	US-07-757-022B-56	Sequence 56, Appl
42	134	2.8	130	4	US-07-757-022B-78	Sequence 78, Appl
43	134	2.8	130	4	US-07-757-022B-86	Sequence 86, Appl
44	134	2.8	138	4	US-07-757-022B-34	Sequence 34, Appl
45	134	2.8	141	4	US-07-757-022B-80	Sequence 80, Appl

## ALIGNMENTS

RESULT 1  
US-09-187-331-6

; Sequence 6, Application US/09187331

; Patent No. 6043056

; GENERAL INFORMATION:

; APPLICANT: Yue, Henry

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Gorgone, Gira A.

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS

; FILE REFERENCE: PF-0631 US

; CURRENT APPLICATION NUMBER: US/09/187,331

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PERL Program

; SEQ ID NO 6

; LENGTH: 873

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: g189650

US-09-187-331-6

Query Match 52.0%; Score 2498.5; DB 3; Length 873;

Best Local Similarity 53.1%; Pred. No. 86-235;

Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

Qy	10	EQPVKE---NTLK---KYKIACIVLLALLVIMSLGLGLGLRLKLEKQSGCRKKCFDAS	62
Db	5	EEPLEKAARARTAKDFTYKYLVLVSVLVTLLGCGIFGLKPKSCAKEVKSGRCFCERT	64
Qy	63	FRGLENCRDVACKDGCWPFEDTCVETRIWVKNKFRGCTRLASLCSDDCLQK	122
Db	65	F---GNCRCDACVGLNCCLDYQETCEPEHIWTCNKRFCGKRLTSLCACSDCKDK	121
Qy	123	KDCADYKSVCCQGETSWLEENCDDTAQQCQCEGFDLPVPVILFSPMDGFRABLYTWDILMP	182
Db	122	GDCCINYSVCCQGEKSWVEEPCESINEPCQAGFETPTLLFSLDGFRAEYLHTWGLLP	181
Qy	183	NINKLTCGIIHSHKVRAMYPTKTFNNHYTIVTGLPESHGIIIDNNMYDNLNKPFLSSK	242
Db	182	VISKLKCCYTKNNRPFVPTKTFNNHYSIVTGLYPESHGIIIDNNMYDNPKNAGSFLSK	241
Qy	243	EQNPPAWHGGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPIYMPYKGVPPFERISTL	302
Db	242	EKENPEWYKGEIWTAKYQGLKSGCTFFWPGSDVEINGIFPDIYKMGVSVPPFERILAV	301
Qy	303	LKWLDLPKERRPRFTYMTFFPEPDSGHAGGPPVSARVIALQVVDHAFQMLMEGLKQRLH	362

Db 302 LQWLQJPKDRPHFYTLYLEEPSSGHSYGVSEVVKALQVRDGMVGMMLDGLKSLNLH 361  
Qy 363 NCVNIIILLADHGMDOQYCNKMEYMTDYPPRINFYMYEGPAPRIAHNIPHPDFFSSE 422  
Db 362 RCLNLIILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGAARLRSDVPDKYSENYEG 421  
Qy 423 IVNLSCKRKPDPKPKVLTDLKRLHYAKNVRIDKVLHFDVQOW-LAVRSKNTNCGG 481  
Db 422 IARNLSCREPNQHPKPKVHLKRLHFAKSDRIEPLTFYLDPMQWQALANPSEKRYCGSG 481  
Qy 482 NHGYNNEFRSMEAFILAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLNH 541  
Db 482 FHGSDNVFSNQALFVGYGPGFKHGEADTFENIEVYNLMCDLLNLTAPNNGTHGSLNH 541  
Qy 542 LKVPYFEPHSHAEVSKFSVCGFANPLPTESLDQFC-PHLQNSTOLEQVNMNLNLTQEEI 600  
Db 542 LKKNPVYTPKPKVHPLVQCPFTRN-PRDNLGCSNPSI---LPTEDEPQTFNLTVAE 597  
Qy 601 TATVKVNLPPGRPRVLQKNVDHCLLYHREVYSGFGKAMPMWSSVYVLPQDGTSPLPPT 660  
Db 598 KLIKHEPLVGRPRVLQKENTICLLSQHFMGSYQDILMPLWTSYTVDR--NDSFSTED 655  
Qy 661 VPCLRADVVRPSESQKSFYLADKNITHGFLYPPASNRTSDSOY-DALITSNLVPMYE 719  
Db 656 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGLSPQLNKNSSGIYSEALLTNIIVPMYQ 715  
Qy 720 EPRKMDYFHSVLLIKHATERGNVNVSGPIFDYNDVGHDFADPDEITKH---LANTDVI 776  
Db 716 SFQVWRYFHDLLRKYAERGNVNVSGPVDFDYDGRCDSENLROKRRVIRNQEILI 775  
Qy 777 PTHYFVLTSCKNKSHTPCNCPGLDVLPIIHPHRTNVESCEPKGPEALWVEERFTHI 836  
Db 776 PTHFFVLTSCKDTSQTLPHCEN-LDTLAFILPHRTDNSESCHGKHDSWVEELLMLHR 834  
Qy 837 AVRDVELLTGLDGYQDKVQPVSEILQKTYLPTF 871  
Db 835 ARITDVEHITGLSFYQQRKEPVSDILKLTLPF 869

## RESULT 2

US-09-470-946-6

; Sequence 6, Application US/09470946

; Patent No. 6358923

; GENERAL INFORMATION:

; APPLICANT: Yue, Henry

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Gorgone, Gina A.

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS

; FILE REFERENCE: PF-0631 US

; CURRENT APPLICATION NUMBER: US/09/470,946

; EARLIER FILING DATE: 1999-12-22

; EARLIER APPLICATION NUMBER: US 09/187,331

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PERL Program

; SEQ ID NO 6

; LENGTH: 873

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: g189650

US-09-470-946-5

Query Match

Best Local Similarity 52.0%; Score 2498.5; DB 4; Length 873;

Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

Qy 10 BOPVKK---NTLK---KYKTIACIVLALLVIMSLGLGLGLKLEKQSCRKKFPDAS 62

Db 5 BEPLEKKAARTAKDPNTYKVLISLVSVCVLTLLTGICIFGLKPSCAKEVSKRCRCFERT 64

```
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 873
; TYPE: PRT
; ORGANISM: H. sapiens
; US-09-438-906-2

Query Match      52.0%; Score 2498.5; DB 4; Length 873;
Best Local Similarity 53.1%; Pred. No. 8e-235;
Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

Qy 10 EQPVKK---NTLK---KYKIACIVLLALLVIMSLGLGLGLKLEKOGSKCKCFDAS 62
Db 5 EEPLKAAARTAKDPNTYKVLVSVLCVLTILGCIIFGLKPSCAKEVSKGRCFERT 64

Qy 63 FRGLENCRCVACDRGDCWDFTCVESTRIWMCNFRCEGTRELEASLCSDDCLOK 122
Db 65 F---GNCRDCAACVBLGNCCLDYQETCIEPEHIWTCNFKRCEKRLTRSLCACSDCKDK 121

Qy 123 KDCADYKSVCOGETSWLENCNTAQOQCEGDFLPVILFSDMGFRAEYLYTWDTLMP 182
Db 122 GDCINYSVCOGEKSWEEPCESINEPQCPAGFETPTLLFSLDGFRAEYLYHTWGLLP 181

Qy 183 NINKLKTGCIHSKYVAMYPKTFFPNHYTIVTGLYPESHGIIIDNNYDVNLKNFSLSK 242
Db 182 VISLKKCGTYTKNRPVPTKTFPNHYSLIVTGLYPESHGIIIDNNYDVNLKNFSLSK 241

Qy 243 EQNNPAWHGQPMWLTAMYQGLKAATYFPGSEVAINSPFPIYMPYNGSVPFEBRISTL 302
Db 242 EKNPEWYKGEPIWTAKYQGLSGTFFWPGSDVEINGIFPDYKMYNGSVPFEBRILAV 301

Qy 303 LKWLDLPKAPRPRFTYMTYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQRLNH 362
Db 302 LQWLQPKDRPHFTYLYLEEPDSSGHSYGPVSEVSEVIKALQVVDGVMGMDGLKRLNH 361

Qy 363 NCNVIILLADHGMQTYCNKMEYMTDYFFRPNFYMEGPAPRAIRAHNIPHDFFSFNSE 422
Db 362 RCLNLIJSDHMEQSGCKYIYLNKYLGVKVIKVIYGAARLRSDVPDKYSEYEG 421

Qy 423 IVNLSCKRQDQHPKPYLTLPDLKRLHYAKVNRIDKHLFVDOQW-LAVRSKNTNCGGG 481
Db 422 IARNLSCEEENQHPKPYLKHFLKRLHFAKSDRIEFTFLYLDQWQALNPSERKYCGSG 481

Qy 482 NHGVNNEFRSMEALFLAHGFSFKEKTEVEPEENIEVYNLMCDLLRIQAPNNGTHGSLNH 541
Db 482 FHGSDNVFSNQALFVGYGFGFKHGEADTFENIEVYNLMCDLLNLTAPNNGTHGSLNH 541

Qy 542 LLKVPFYPGSHAEVSKESVCGFANPLTBSLDFC-PhLQNSTQLEQVNMNLTOBEI 600
Db 542 LLKNPVYTPKXPEVHPLVQCPETRN-PRDNLGSCNPSI---LPIDFQTNLTVAEE 597

Qy 601 TATVKVNLPGRRPRVLQKNDHCLLYHREYVSGFGKAMPWSSYTVPLGDTSPLPPT 660
Db 598 KIHKETLPYGRPRVLQKNTICLLSQHFMGSGISQDILMPLMTSTVTDK--NDSPSTD 655

Qy 661 VPCDLRADVRPPSESKCSFYLAADKNIITHGFIYPPASNRITSQY-DALITSNLVPMYE 719
Db 656 FSNCLYQDFRPLSPVHKCSFYKNNTKVSXGFLSPQNLKNSSGIYSEALLTNIIVPMYQ 715

Qy 720 EFRKQWYFHSVLIIKHATERNGVNVVSGIFDYNVDGHFADDEITKH---LANTDVPI 776
Db 716 SFQVIRYFHTLLARKAERNGVNVVSGIFDYNVDGHFADDEITKH---LANTDVPI 775

Qy 777 PTHFVVLVTSCKNSHPENCGLVDLPTIIPRPTNVSCPEGKPEALWVEERTTAHI 836
Db 776 PTHFVVLVTSCKNSHPENCGLVDLPTIIPRPTNVSCPEGKPEALWVEERTTAHI 834

Qy 837 ARVEDVELLTGLDFYQDKVQVPSIQLKTLPTFF 871
Db 835 ARITDVEHITGLSFYQKQKPEVSDILKTLPTFF 869
```

```
RESULT 4
US-08-392-946-1
; Sequence 1, Application US/08392946
; Patent No. 5939269
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: The Regents of the University of California
; APPLICANT: Goldfine, Ira D.
; APPLICANT: Grupe, Andrew
; APPLICANT: Maddux, Betty A.
; APPLICANT: Spencer, Steven
; APPLICANT: Stewart, Timothy A.
; TITLE OF INVENTION: Antagonists to Insulin Receptor Tyrosine
; TITLE OF INVENTION: Kinase Inhibitor
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,946
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/182241
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0875P1PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-392-946-1

Query Match      52.0%; Score 2498.5; DB 2; Length 925;
Best Local Similarity 53.1%; Pred. No. 8.8e-235;
Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

Qy 10 EQPVKK---NTLK---KYKIACIVLLALLVIMSLGLGLGLKLEKOGSKCKCFDAS 62
Db 57 EEPLKAAARTAKDPNTYKVLVSVLCVLTILGCIIFGLKPSCAKEVSKGRCFERT 116

Qy 63 FRGLENCRCVACDRGDCWDFTCVESTRIWMCNFRCEGTRELEASLCSDDCLOK 122
Db 117 F---GNCRDCAACVBLGNCCLDYQETCIEPEHIWTCNFKRCEKRLTRSLCACSDCKDK 173

Qy 123 KDCADYKSVCOGETSWLENCNTAQOQCEGDFLPVILFSDMGFRAEYLYTWDTLMP 182
Db 174 GDCINYSVCOGEKSWEEPCESINEPQCPAGFETPTLLFSLDGFRAEYLYHTWGLLP 233

Qy 183 NINKLKTGCIHSKYVAMYPKTFFPNHYTIVTGLYPESHGIIIDNNYDVNLKNFSLSK 242
Db 234 VISLKKCGTYTKNRPVPTKTFPNHYSLIVTGLYPESHGIIIDNNYDVNLKNFSLSK 293

Qy 243 EQNNPAWHGQPMWLTAMYQGLKAATYFPGSEVAINSPFPIYMPYNGSVPFEBRISTL 302
Db 294 EKNPEWYKGEPIWTAKYQGLSGTFFWPGSDVEINGIFPDYKMYNGSVPFEBRILAV 353

Qy 303 LKWLDLPKAPRPRFTYMTYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQRLNH 362
```

[illegible]

## RESULT 5

US-08-504-169-1  
; Sequence 1, Application US/08504169  
; Patent No. 5968508  
; GENERAL INFORMATION:  
; APPLICANT: Goldfine, Ira  
; APPLICANT: Grupe, Andrew  
; APPLICANT: Henzel, William  
; APPLICANT: Maddox, Betty  
; APPLICANT: Spencer, Steven  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Harrison, Denise M.  
; TITLE OF INVENTION: Antibodies to Insulin Receptor Tyrosine Kinase Activation Inhibitors  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/504,169  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/14893  
; FILING DATE: 28-Dec-1994

QY 777 PTHYFVVLTSCKNKSHTPENCQWGLDVLPIIHPHRTNVSCEPGKEALWVEERFAHI 836  
 DB 828 PTHFFIVLTSCDTSQTPHLCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELMLHR 886  
 QY 837 ARVRDVELLTGLDFYQDKVQPVSEIILQKTYLPTF 871  
 DB 887 ARITDVEHITGLSFYQORKEPVSDILKXTHLPTF 921

## RESULT 6

PCT-US94-14893-1  
 ; Sequence 1, Application PC/TUS9414893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; APPLICANT: The Regents of the University of California  
 ; APPLICANT: Goldfine, Ira D.  
 ; APPLICANT: Grupe, Andrew  
 ; APPLICANT: Maddux, Betty A.  
 ; APPLICANT: Spencer, Steven  
 ; APPLICANT: Stewart, Timothy A.  
 ; TITLE OF INVENTION: Antagonists to Insulin Receptor Tyrosine Kinase Inhibitor  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/14893

; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/182241  
 ; FILING DATE: 14-JAN-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kubinec, Jeffrey S.  
 ; REGISTRATION NUMBER: 36,575  
 ; REFERENCE/DOCKET NUMBER: 875P1PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE:  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 925 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear

PCT-US94-14893-1

Query Match 52.0%; Score 2498.5; DB 5; Length 925;  
 Best Local Similarity 53.1%; Pred. No. 8.8e-235;  
 Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;  
 QY 10 EQPVKK---NTLK---KYIACITVIALLVIMSLGLGLGLKLEKOGSCRKCKCFDAS 62  
 DB 57 EEPLKKAARATADPNYKVLVSLVSVLTITLGCIFGLKPSKAEKVKSGRCRFT 116  
 QY 63 FRGLENCRCDVACDRGDCCWDFEDTCTVSTRIMCNKFRCGFTRLEASLCSCDDCLOK 122  
 DB 117 F---GNCRCDAAVGLNCCLDYQETCTIEPEHIWTCNKPFCGEKRLTRSLCACSDCKDX 173  
 QY 123 KCCADYKSVCGGTSTWLENCNTAQSCPEGEFDLPVILFSDGCFRAEYLTYWDTLMP 182  
 DB 174 GDCCINYSVVCQGEKSWBPCSEINBPQCFAGFETPTLLFSLDGFRAEYLHTWGLLP 233

QY 183 NINKLTKCGIHSKMEAMYPYTKTFPNHYTIVTGLYPESHGIDNNMYDVNLKNFSLSK 242  
 DB 234 VISKLKCCGTYTQWRPVYPTTKTFPNHYISIVTGLYPESHGIDNNMYDPKNASFSLSK 293  
 QY 243 EQNNPAWMEGQPMWLTAMTQGLKAATYFWPGSEVAINSGFSPSIYMPYNGSVPPFERISTL 302  
 DB 294 EKENPEWYKGEPIWWTAKYQGLKSGTFFWPGSDVEINGIFPDYIWKYNGSVPPFERILAV 353  
 QY 303 LKWLDLPKAEPRFYTMPEEPDSSGHAGGPPVSAVIALQVVDHAFQMLGKQRLVH 362  
 DB 354 LOWLQPKDERPHERFYTLLEEDSSGHSYGPVSSSVIKALQKQVDMGVMGLMDGLKELNLH 413  
 QY 363 NCVNIILLADHGMDOQYCNKMEYMTDYFPRINFFYMWYEGPAPRIAHNIPHDFFSFNSEE 422  
 DB 414 RCLNLILSDHGMEOGCKYIYLNKYLGDVKNIKVIYGPAAARLPSPDVPKYISFYNEG 473  
 QY 423 IVNLSCRKPDQHFYLPYLPDLPKGLHYAKNVRIDKVLHFDVQOQW-LAVRSKSNVTCGGG 481  
 DB 474 IARNLSCREPNQHFKYLPKHLFAKSDRIEPLTFYLDPOWQALALNFSERKYCSG 533  
 QY 482 NHGYNNEFRSMAEFLAHGSPFKEKTEVEPPFENIYVYMLMCDLLRIQAPANNTHGSLNH 541  
 DB 534 FHGSDNVFSNQALFVGYGPGFKGIEADTFENIEVYMLMCDLLNLTPAPNNGTHGSLNH 593  
 QY 542 LLKVPYEPESHABEVSKFVCGFANPLPTESLDCTC-PHLQNSTOLEQVQNMNLTOEEI 600  
 DB 594 LLKNPVYTPKPKVHPLVQCPFTRN-PEDNLGSCNFSI---LPIDFQTFQNLTVAAE 649  
 QY 601 TATVKVNLPGRPVRLQKNVDHCLLYHREYVSGFKAMRPMWSSYTPQLCDTSPLEPT 660  
 DB 650 KIHKETLPYGRPVQLKENTICLLSQHFMGSGYQDILMPLWTSYTVDR--NDSFSTED 707  
 QY 661 VPDCLRADVRVPPSPESQKSFYLADKNITHGFLYPPASNRSDSOY-DALTSNLVPMYE 719  
 DB 708 FSNCLYQDFRIPLSPVHKCSFYKNTKVSYGFLSPQPLNKNSSGYSALLTNTVPMYQ 767  
 QY 720 EFRKMDYDHSVLLIKHATERGVNVVSGPIEDYVNDGHFDAPDEITKH---LANTDVEI 776  
 DB 768 SFQVIWRYFHDLLRYAERGVNVVSGPVDFDYDGCDSLENLKQKRVIRNQEILI 827  
 QY 777 PTHYFVVLTSCKNKSHTPENCQWGLDVLPIIHPHRTNVSCEPGKEALWVEERFAHI 836  
 DB 828 PTHFFIVLTSCDTSQTPHLCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELMLHR 886  
 QY 837 ARVRDVELLTGLDFYQDKVQPVSEIILQKTYLPTF 871  
 DB 887 ARITDVEHITGLSFYQORKEPVSDILKXTHLPTF 921

## RESULT 7

US-09-438-906-4  
 ; Sequence 4, Application US/09438906  
 ; Patent No. 6465185  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldfine, Ira  
 ; APPLICANT: Trischitta, Vincenzo  
 ; APPLICANT: Pizzuti, Antonio  
 ; APPLICANT: Vigneri, Riccardo  
 ; APPLICANT: Frittitta, Lucia  
 ; TITLE OF INVENTION: Polymorphic Human PC-1 Sequences  
 ; FILE REFERENCE: 9076-089  
 ; CURRENT APPLICATION NUMBER: US/09/438,906  
 ; CURRENT FILING DATE: 1999-11-18  
 ; PRIOR APPLICATION NUMBER: 60/108,853  
 ; PRIOR FILING DATE: 1998-11-18  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 873  
 ; TYPE: PRT  
 ; ORGANISM: H. sapiens  
 ; US-09-438-906-4





QY	304	KWLDPKAEPRFYTMFEEDSSGHAGGPVSA-----	336
DB	291	RWLTLTPDHERSPVAFVSEQDFSGHKYGFPGPEESYGSFPTPAKPKRKVAPKEROER	350
QY	337	-----RVIKALQVVDHAFGLMEGLKORNLHNCVMILIA	371
DB	351	PVAPPKRRRKIHRMDHYAAETRODKVTNPLREIDKIVGQLMDGLAQXLRLRCVNVFVG	410
QY	372	DHGMQDTYCNKMEYMTDYPRINFYMYEGPAPRIR---AHNIPHPDFSFNSEIYRNL	428
DB	411	DHGMEDVTCDTEFSLNSLYLTNVDDITLVPOTGLGRKSKFSNNAKYD-----PKAIANLT	465
QY	429	CRKPDQHPKPYLTDLPKRLHYAKNVRIDKVLHVFVQQW-----LAVRSKNTNC-GGG	481
DB	466	CKKPDQHPKPYLKQHLPKRLHYANNRIEDITHLLVERHHVARKPLDVYKPKSGCKFFOG	525
QY	482	NHGYNNERSNEATFLAHGSPFKETKEVEFENIEVYMLCDLLRTOCAPNNGTHGSLNH	541
DB	526	DHGFDPNKVNSQMTVFGVGPTFKYTKVPFENIELYNNVNCDDLGLUKPAPNNGTHGSLNH	585
QY	542	LKVPFYEPFSHAVERSFVSGFANPLPTELSLDFC-FHLQNSTQLEQVNMNLNQBEI	600
DB	586	LLRNTNTPRTPPEEVRPNYEGIMYLOSDFDLGCCTDKKVEPKXKULDELKGLHTK----	641
QY	601	TATVAVNLPGRPVRLQNVGDHCLLYHREYVSGFKAMRPMWSSYTVPLQDGTSLPPT	660
DB	642	GSTERHLLYGRPAVLNR-TRYDLYLTDTSFSGSEIFLMLLTSTVSKQAEVSSVDH	700
QY	661	VPDCURADVRRPPSESQKCSFYLDKNITHGFLYPASNRTSDSYDALITNLVPMYEE	720
DB	701	LTSQVRPDRVSPSPSQNLAYKNDKQMSYGFLLPPVYLSSEPAAKYDAFLVTNMVMPYA	760
QY	721	FRKMDWYFHSVLLIKHATERGVNVVSGPFDVNYDGHFDAPDEIKHLANTDVPIPTHY	780
DB	761	FRKWNVYQFRLVKYASERGVNVISGPFDDYDGLHDEDKIKQYVEGSSIPVPTHY	820
QY	781	FVYLTSCKNGHTPENCPSGLVDLFPFIHPRTNVESCPEGKPEALWEEFTHAIAVR	840
DB	821	YSIIITSLDFTQADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWSELMKMHRTARVR	880
QY	841	DVELLTGLDGFQDKVQPVSEILOKTYLPTFEITI	875
DB	881	DIEHTSLDFFRKTSRSPYELTUKTYLHTYSEI	915
RESULT 9			
US-08-977-221-69			
; Sequence 69, Application US/08977221			
; Patent No. 6084069			
; GENERAL INFORMATION:			
; APPLICANT: UNITED STATES OF AMERICA; DEPT.			
; APPLICANT: OF HEALTH AND HUMAN SERVICES			
; TITLE OF INVENTION: MOTILITY STIMULATING			
; TITLE OF INVENTION: PROTEIN STIMULATING			
; TITLE OF INVENTION: THERAPY			
; NUMBER OF SEQUENCES: 69			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: MORGAN & FINNEGAN			
; STREET: 345 PARK AVENUE			
; CITY: NEW YORK			
; STATE: NEW YORK			
; COUNTRY: U.S.A.			
; ZIP: 10154			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy Disk			
; COMPUTER: IBM PC Compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: WordPerfect 5.1			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/977,221			
; FILING DATE:			
; CLASSIFICATION:			

Db 526 DHGPDNKNVSMQTVFVGVPTFKYKTKVPPFENIELVNMWCDLLGLKAPNNGTHGSLNH 585  
Qy 542 LLKVPFFPEPSHAEVSVKESVCGFANPLPTESLQFC-PHLQNSTQLEQVNMNLTOBEI 600  
Db 586 LLRTNTRFTMPSEVTRPNVPGIMYLOSDFDLGCTDDKVEPKKDUKLNKLHTK---- 641  
Qy 601 TATVKNLPPFRPRVLQKNDHCLLYHREYVSGFGKAMRMPMWSSYTVPLQDGTSPPLPT 660  
Db 642 GSTERHLLYGRPAVLYR-TRYDILYHTDFESGVSEIFLMLLWTSYTVSKQAEVSSVDPH 700  
Qy 661 VPCDLRADVRVPSSEKSCFYLADKNITGFLYPPASNRSTDSQDALITSNLVPMEYEE 720  
Db 701 LTSCVRPDRVRSFQSNCLAYKNDQMSYGLFPPYLSSSPEAKYDAFLVNMVMPYPA 760  
Qy 721 FRKWDYPHSVLLIKHATERNGVNVSGPIFDYNDGHFDAPDEITKHLANTDVPIPTHY 780  
Db 761 FKRVMNYFORVLVKYASERNGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIFVPTHY 820  
Qy 781 FVLTSCNKSHTPENCPCGMLDVLFPFIIPHRPTNVSCPEGKPEALWVEERFTHARVR 840  
Db 821 YSIITSCDLFTQPADKCDGFLSVSSFLPHRPDNESSCNSEDESKWVELMKQHTARVR 880  
Qy 841 DVELLTGLDFYQDKVQVSEILOLKYLPFTFTTI 875  
Db 881 DIEHLTSLDFPRKTSRSPYELTLTKYLHTVSEI 915

## RESULT 10

US-09-483-831B-69  
; Sequence 69, Application US/09483831B  
; Patent No. 6417338  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY  
; APPLICANT: LIOTTA, LANCE  
; APPLICANT: SCHIFFMANN, ELLIOTT  
; APPLICANT: KRUTZCH, HENRY  
; APPLICANT: MURATA, JUN  
; TITLE OF INVENTION: AUTOTAXIN: MOTILITY STIMULATING PROTEIN USEFUL IN  
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2026-4149US4  
; CURRENT APPLICATION NUMBER: US/09/483,831B  
; CURRENT FILING DATE: 2000-01-17  
; PRIOR APPLICATION NUMBER: 07/822,043  
; PRIOR FILING DATE: 1992-01-17  
; PRIOR APPLICATION NUMBER: 08/249,182  
; PRIOR FILING DATE: 1994-05-25  
; PRIOR APPLICATION NUMBER: 08/346,455  
; PRIOR FILING DATE: 1994-11-28  
; PRIOR APPLICATION NUMBER: 08/977,221  
; PRIOR FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 69  
; LENGTH: 915  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Polypeptide  
US-09-483-831B-69

Query Match 42.5%; Score 2040.5; DB 4; Length 915;  
Best Local Similarity 40.5%; Pred. No. 5.2e-190;  
Matches 379; Conservative 177; Mismatches 268; Indels 111; Gaps 11;  
Qy 24 ACIVLLALLVMSLGLGGLGRKLEK-----QGSCKKCFDASF 63  
Db 9 SCIIISLFTFVAGVSCIGTAKIRKAEWGEGPTVLSDSPWNISCKGRCFELQE 68  
Qy 64 RGLNCRCDVACDRGDCWDFEDTCVVESTRIWMCKNFKRGETRLFASLCSDCLQKX 123  
Db 69 AGPPDCRCDNLKSYTSCCHDFDLCLKTARGWECTKDRCGEVNENACHGSDCLARG 128

Qy 124 DCCADYKSCQGETSLEBNCDTAQOSQCEGFDLPPLVILFMSDMGFAEYLYTWDTLMPN 183  
Db 129 DCCNTQVQVCKGESHVWDDDEIKAACECPAGFVPPILIIISVDGFRASYMKGSKVMPN 188  
Qy 184 INKLKTCGHSKYMAMYPKTFTFNNHYITVGLYFESHGIIDNNMYDNLNKNFSLSKSE 243  
Db 189 IEKLSRSGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDFVDFATHLRGRE 248  
Qy 244 QNPPAWHGGOPMMLTAMYOGKAAATVFWPGSEVAINGSPSPSYMPYNGSVFPEERISTLL 303  
Db 249 KFHRWGWGOLPMTATKQGVKAGTFW-----SVVIFH-----ERILLTIL 290  
Qy 304 KWLIDLKABRPPRYTYWFEEDSSGHAGGPVSA----- 336  
Db 291 RWLTLPHDERPSVYAFYSBQDFSGHKYGFPGFEESYSGSPFPAKRPKRVAPKRRQER 350  
Qy 337 -----RVIKALQVVDHAFGLMEGLKORNLHNCVNIILLA 371  
Db 351 PVAPPKRRKRKHHRMDEHYAAETROQKMTNPLREIDKIVGQIMDGLKQLKLRQNVVIFVG 410  
Qy 372 DHGMDQTYCNKMEYMTDYFPRINFFYMEGAPAPRIR--AHNIPHDFFSFSEIIVRNLS 428  
Db 411 DHGMDVTCDRTEFLSNLYLTNVDITLVEGTIGRIRSKFSNNAKYD-----PKAIANLT 465  
Qy 429 CRKPDQHFKPYLTPDLPKRLHYAKXVRIDKVLHFVDOOM-----LAVRSKSNNTNC-GGG 481  
Db 466 CKPDDQHFKPYLQKLPKRLHYANNRIED IHLVERRHVVARKDLYVKSPGKCPFGQ 525  
Qy 482 NHGYNNEFRSMEAI FLAHGSPKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLNH 541  
Db 526 DHGFDNKNVSMQTVFVGVPTFKYKTKVPPFENIELVNMWCDLLGLKAPNNGTHGSLNH 585  
Qy 542 LLKVPFFPEPSHAEVSVKESVCGFANPLPTESLQFC-PHLQNSTQLEQVNMNLTOBEI 600  
Db 586 LLRTNTRFTMPSEVTRPNVPGIMYLOSDFDLGCTDDKVEPKKDLNKLHTK---- 641  
Qy 601 TATVKNLPPFRPRVLQKNDHCLLYHREYVSGFGKAMRMPMWSSYTVPLQDGTSPPLPT 660  
Db 642 GSTERHLLYGRPAVLYR-TRYDILYHTDFESGVSEIFLMLLWTSYTVSKQAEVSSVDPH 700  
Qy 661 VPCDLRADVRVPSSEKSCFYLADKNITGFLYPPASNRSTDSQDALITSNLVPMEYEE 720  
Db 701 LTSCVRPDRVRSFQSNCLAYKNDQMSYGLFPPYLSSSPEAKYDAFLVNMVMPYPA 760  
Qy 721 FRKWDYPHSVLLIKHATERNGVNVSGPIFDYNDGHFDAPDEITKHLANTDVPIPTHY 780  
Db 761 FKRVMNYFORVLVKYASERNGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIFVPTHY 820  
Qy 781 FVLTSCNKSHTPENCPCGMLDVLFPFIIPHRPTNVSCPEGKPEALWVEERFTHARVR 840  
Db 821 YSIITSCDLFTQPADKCDGFLSVSSFLPHRPDNESSCNSEDESKWVELMKQHTARVR 880  
Qy 841 DVELLTGLDFYQDKVQVSEILOLKYLPFTFTTI 875  
Db 881 DIEHLTSLDFPRKTSRSPYELTLTKYLHTVSEI 915

## RESULT 11

PCT-US95-06613-69  
; Sequence 69, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK

COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. JUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 915  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: A2058 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PCT-US95-06613-69

Query Match 42.5%; Score 2040.5; DB 5; Length 915;  
Best Local Similarity 40.5%; Pred. No. 5.2e-130;  
Matches 379; Conservative 177; Mismatches 268; Indels 111; Gaps 11;

QY 24 ACIVLLALLVMSLGLGLGLRLEK-----QSGCRKCKPDASF 63  
DB 9 SQIISLTFVAVGVSICIGFTHAKRAEGWEGPPTVLSDSPWNTISSGCKGRCELOE 68  
QY 64 RGLNCRDVAKDRGDCWDFCTVSTIWNCKNFRGETLEASLCSDDCLQKX 123  
DB 69 AGPPDCRDLCKSYTSCCHDFDELCLATARGWECTKDRCEVRNEENACHCEDCLARG 128  
QY 124 DCADYKVCOCETSWLENCDDTAQOCPSGFDLPPVILFSDMDGRAEYLYTWDILMPN 183  
DB 129 DCTNYQVCKESHVDDCEEIKAAACPCAGFVRPLIFSDVDFRASYKKGSKWMPN 198  
QY 184 INKLKTCGHSKYMAYPTKTFPNHYITVGLYPESHGIIIDNNMVDNKLKFSLSKE 243  
DB 189 IEKLARSCGTHSPYRFPVPTKTFPNLYLATGLYPESHGIVGNSYDVFEDATFHLGRE 248  
QY 244 QNNPAWHGQPMWLTAMVQGLKAATYFWPGSEVALNGSPSIYMPYNGSVFFPEERSTLL 303  
DB 249 KFNHRWGGQPIWITATQGVKAGTFW-----SVVIPH-----ERRILITL 290  
QY 304 KMLDLPKARPRFYTWYFEEPDSSGHAGPVSA-----336  
DB 291 RMLTLPDHERPSVAFYSEQDPDSGHKYGFPGPBESSYSGSPFTPAKRPKRVAPKRRQER 350  
QY 337 -----RVIKALQVVDHAFGLMEGLKORNLHNCVNIILIA 371  
DB 351 PVAPPKGRARKHRMDHYAAETRODKMNPLREDKIVGQLMDGLKQLKLRVCUNVIFVG 410

QY 372 DHGMDQTYCNKMYMTDYPPRAINFYMYEGPAPRIR---AHNIPHDFFSFNSEIURNLS 428  
DB 411 DHGMDVTCDRTEBFLSNLYLNVDDITLVPCTLGRIRSKFSNNAKYD-----PRAITANLT 465  
QY 429 CRKPDQHFKPYPDLPKRLHYAKNVRIDKVHLFVDDQM-----LAVRSKSNINC-GGG 481  
DB 466 CKKPDQHFKPVLQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKPSGKCFQ 525  
QY 482 NHGYNNEFRSKEALFLAHGFSFKETVEBPENIEVYNLMCDLLRQAPNNGTHGSLAH 541  
DB 526 DHGFDNKVMSQTVFVGYGTFKTKVPPFENIELYNVWCDLLGKPKAPNNGTHGSLNH 585  
QY 542 LLKVPFYPEPSHAEVSKFSYCGFANPLPTESLDFCF- PHLQNSTOLEQVNMQLNLQEEI 600  
DB 596 LLRTNTRPTWPEBETRPVPGIMYLQSDPDLCTCCKVPEKPKLDELNKRHLTK- --- 641  
QY 601 TATVKNLPGRPVYLOKNDHCLLYHREYVSGFKAMRPMWSSYTVQLGDTSLPPT 660  
DB 642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESYGSIFLMLLWTSYTVSKQAEVSVPOH 700  
QY 661 VPDCLRADVRVPSESQKCSFYLLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720  
DB 701 LTSCVRPDRVSPFSQNCCLAYKNDKQMSYGLFPFVLSSSPEAKYDAFLVTNNVMEYFA 760  
QY 721 FRKXMDYFHSVLILKHATERNGVNVSGPIFDYNDGHEPADDEITKHLANTDVPIPTHY 780  
DB 761 FRVWNYFORVLVKKYASERNGVNLSGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHY 820  
QY 781 FVVLTSCKNKSHTPENCFCGMLDVLPTIIPRPTNVESCPGKPEALWBERFTHAIVARV 840  
DB 821 YSIIISCLDFTQPADKCDGFLSVSSFLPHRPDNESSCNSSEDESKWBEELMQHTARVR 880  
QY 841 DVELLTGLDFYQDKVQPVSVSILQIKYLPFTFETI 875  
DB 881 DIEHLTSLDFRKTSRSPYELITLKYVLYHTYSEI 915

RESULT 12  
US-08-346-455B-34  
Sequence 34; Application US/08346455B  
Patent No. 5731167  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,455B  
FILING DATE: 28-NOV-1994  
CLASSIFICATION: 530  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043

FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149PCT  
TELEPHONE: (212) 751-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 829  
TYPE: amino acid  
STRADEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: Melanoma  
CELL LINE: A2058  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: Putative protein  
OTHER INFORMATION: sequence of A2058 Autotaxin  
US-08-346-455B-34

Query Match 40.9%; Score 1964.5; DB:1; Length 829;  
Best Local Similarity 42.4%; Pred. No. 1.2e-182;  
Matches 363; Conservative 164; Mismatches 239; Indels 91; Gaps 10;

QY 82 CWDPEDTVCSETRIMCKNFKRCGTRLEASLCSDCLOKCCADYKSCVCOGETSWLE 141  
DB 1 CHDFELCLTKARGWECTKDRGVEERNEACHSEDCCLARGDCCTNYQVCKGESHWD 60  
QY 142 ENCDTAQSQSQCEGFDLPVILFSDMGFERAYLYTWTMPNINKLTGCIHISKYRMY 201  
DB 61 DCESEIKAAECAPAGVRPLIIFSDGFRASYMKGKSKWENIEKLRSCGTHSPYMRPY 120  
QY 202 PKTPEPNHYTIVTGLYPESHGIIIDNNMYDVNMLKNFSLSSKEQNPANWHGQPMWLTAMY 261  
DB 121 PKTPEPNLYTLATGLYPESHGIVGNSMYDVPVDFATFHLRGREKFNHRWGGQPLWITATK 180  
QY 262 QGLKAATYFWGSEVAINGSPSPYIMYNGSVPEERISILLKWLDPKAEERFRFTWTF 321  
DB 181 QGVKAGTFW-----SVVIPH-----ERRILTLRLWTLDPHERPSVAFYS 222  
QY 322 BEPDSSGHAGPVSA----- 336  
DB 223 EQPDSGHKYGFPGESESSYSPTPAKRPRKVPKROERFVAPPKERRKIHMDHY 282  
QY 337 -----RVIALQVDFHAFGLMEGLKORNLHNCVNIILLADHGDQTYCNKMEYMTDY 389  
DB 283 AAEFTQDKXTNPLREIDKIVQQLMDGLKRLRCVNVIFVGHGMEVDTCDSFELSNY 342  
QY 390 FPRINFFYMGAPAPRIR--AHNIPHDFFSSEIEVRNLSCKRPDQHFKPVLTPDLPK 446  
DB 343 LTNVDDITLVPTGLRIRSKESNAKYD-----PKAIANLTCKKPDQHFKPVLKQHLPK 397  
QY 447 RLHYAKNVRIDKIVHLFYDQW-----LAVRSKNTNC-CGGNHGYNFRSNEALFLAH 499  
DB 398 RLHYANNRIRIDHLLVRRHWHARKPLDVYKPSGKCFQGDHGEDNKVNSQMTVFVGY 457  
QY 500 GPSFKETEVFFENIEVYNLMCDLLRIQAPNNGTHGSLMHLKVPFYPPEPSHAEVSXF 559  
DB 458 GPTFKYKTKVPPFENIELYNVCMCLGLKAPNNGTHGSLMHLRTNTFRPTMPEEVTRP 517

QY 560 SVCGFANPLPTESLDCEFC-PHLQNSTGLEQVQNMALNLTQEEITATVKNLPGRRPVLOK 618  
DB 518 NYPGIMYLOSDDDIGCTCDKVKPKKLDLNLRLHTK-----GSTERHLLYGRPAVLRY 573  
QY 619 NVDHCLLYHREYVSGFGKAMRMWMSYTPVPLGDTSPPLPPTVPCDLRADVRVPSESO 678  
DB 574 -TRYDILYHTDFESGYSEIEFLMLLWTSYVSKQAEVSSVPDHLTSCVRPDVRSFSON 632  
QY 679 CSFYLADKNITHGPLYPPASNRISDSQYDALITSNLVPMYSEFRKMWDFHSLVLLKHT 738  
DB 633 CLAYKNDKQSYGFLFPYLSUSSPEAKYDAFLVNNVMTYPAFRVNNYFORVLVKKYS 692  
QY 739 ERGNVNVSGPIFDYNDYDGHFADPEITIKHLANTDVPITPHYFVVLTSCKNSHTPENCP 798  
DB 693 ERGNVNVSGPIFDYDYLHDTEDKIKQVVESSIPVPHYVSIITSCLDFTQPADKCD 752  
QY 799 GWLDVLPILPHRPTNVESCEPKPEALWVEERFTHAIAVRDVELLTGLDFYQDKVQPV 858  
DB 753 GPLSVSSFILPHRPDNEESSNSESSESKWVEELMQLQHTARVDRDIEHLTSLDFFRKT 812  
QY 859 SEILQLKTYLPTFETII 875  
DB 813 PEILTKLYLHYESEI 829

RESULT 13  
US-08-977-221-34  
; Sequence 34, Application US/08977221  
; Patent No. 6084069  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA: DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,221  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; APPLICATION DATA: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 751-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 829  
; TYPE: amino acid



Db 283 AAEATQDKMTNPLREIDKIVGQMDGLKQLKLRRCVNVIVFGDHGMDVTCDRTEFLSN 342  
Qy 390 FPRINFFMYEGAPRIR---AHNIPDFFSFNEEIVRNLSCEKDPDHPKPYLTDLPEK 446  
Db 343 LTNVDDITLVPGLGRISKSFNNAKYD-----PKAIANLTCKKDPDHPKPYLKQHLPK 397  
Qy 447 RLHYAKNVRIDKVLHFDVQW-----LAVRSKNTNC-GGGNHGYNNEFRSMEALFLAH 499  
Db 398 RLHYANNRRIEDIHLVRRWHVARKPLDVVKKPSGKCFQGDHGFNDKNSMOTVFVGY 457  
Qy 500 GPSFKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLNHLKYPFVEPSSHAEVSKF 559  
Db 458 GTTFYKTKVPPFENIELYNMCDLLGLKAPNNGTHGSLNHLRTNTPFTMPEEIVRP 517  
Qy 560 SVCGFANPLFTESLDFCF-PHLQNSTOLEQVQNMNLNTOBEITATVKVNLFPGRPRVLQK 618  
Db 518 NYPGIMYLSQDDDLGCTCDDVPEPKNLDELNKLTKK---GSTBERHLLYGRPAVLYR 573  
Qy 619 NVDHCLLYHEEYVSGFGKAMPMWSSYTVPLQDGTSPDPTVDDCLRADVRVPPSESOK 678  
Db 574 -TRYDILVHTDPSGYSIFLMLNTSYTSVKQEVSVDPDLTSCURPDVRVSPFSQ 632  
Qy 679 CSFYLDXNTHGFLYPPASNRTSDSQYDALITSNLYMYEERKWDYFHSVLLIKHAT 738  
Db 633 CLAYKNDXQMSYGLFPYLSSEPAKYDAFLVTNMYVPAKRVWNYFORVLKVKYAS 692  
Qy 739 ERNGVNVSGPIFDYNDGHFADPEITKHLANTDVPITHYFVLTCKNKSHTPENC 798  
Db 693 ERNGVNVISGPIFDYDGLHDTEDKIQYVEGSSIVPTHYYSIITSCLDFTQADKCD 752  
Qy 799 GMLDVLPIIHRPTNVESCEGPBALWVEERTAHIAVRVVELLTGLDFTYQDKVQV 858  
Db 753 GLSVSSFILPHRPNEESSNSEDSEKWEELMOMHTARVDIEHLTSLDFRKTSSY 812  
Qy 859 SEILOKTYLPTFTTI 875  
Db 813 PEILTKYLTHTYSEI 829

## RESULT 15

PCT-US95-06613-34

Sequence 34, Application PC/TUS9506613

GENERAL INFORMATION:

APPLICANT: STRACKE, MARY; LIOTTA, LANCE;

APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,

APPLICANT: HENRY; MURATA, JUN

TITLE OF INVENTION: MOTILITY STIMULATING

TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESS: MORGAN &amp; FINNEGAN

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: U.S.A.

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06613

FILING DATE: 24-MAY-1995

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/346,455

FILING DATE: 28-NOV-1994

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/249,182

FILING DATE: 25-MAY-1994

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 829  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: Melanoma  
CELL LINE: A2058  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: Putative protein  
OTHER INFORMATION: sequence of A2058 Autotaxin  
PCT-US95-06613-34

Query Match

Best Local Similarity 40.9%; Score 1964.5; DB 5; Length 829;

Matches 363; Conservative 164; Mismatches 239; Indels 91; Gaps 10;

Qy 82 CWFEDTCVSTRWVKNKPRCGFTRLEASLCSDDLQKQCCADYKSVCGEISWLE 141  
Db 1 CHDFDELCLKTARGWECTKRCGEVRNEENACHSEDCLAGDCCTNYQVVCGESHW 60  
Qy 142 ENCDTAQSQSCPEGFDLPVILFSDMGPRARYLTWTLMPTNKTKTCGTHSKYWE 201  
Db 61 DCEEBKAAECAPAGFVRPLIFSVDFRASYMKKSKWMPNIEKLSCTGTHSPYR 120  
Qy 202 PTKTFPNHYITVGLYPESHGIIIDNNYDNLNKNFSLSSKEQNNPAWVGQPMWLT 261  
Db 121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWVGQ 180  
Qy 262 QGLKAATYFEPGSEVAINGSPSPYMPYNGSVPPPEERISTLLKWLDPKAEPR 321  
Db 181 QGVKAGTFW-----SVVIPH-----ERRILILRWLTLPDHERSVAFYS 222  
Qy 322 EEPDSSCHAGGPVSA-----SVVIPH-----ERRILILRWLTLPDHERSVAFYS 336  
Db 223 EQPDFSGHKYGFPGPESSYSGPFTPAKPKRVAPKQRQERPVAPPKGRRIHRMD 282  
Qy 337 -----RVIKALQVVDHAFGLMEGLKQRLNHCNVIILADHGMDOQTCNRYEY 389  
Db 283 AAEATQDKMTNPLREIDKIVGQMDGLKQLKLRRCVNVIVFGDHGMDVTCDR 342  
Qy 390 FPRINFFMYEGAPRIR---AHNIPDFFSFNEEIVRNLSCEKDPDHPKPYLTDL 446  
Db 343 LTNVDDITLVPGLGRISKSFNNAKYD-----PKAIANLTCKKDPDHPKPYL 397  
Qy 447 RLHYAKNVRIDKVLHFDVQW-----LAVRSKNTNC-GGGNHGYNNEFRSME 499  
Db 398 RLHYANNRRIEDIHLVRRWHVARKPLDVVKKPSGKCFQGDHGFNDKNSMOT 457  
Qy 500 GPSFKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLNHLKYPFVEPSSH 559



Db 458 GPTFKYKTKVPPFENIELYNNVCDLLGLKPAPNNGTHGSLNELLRTNTRPTMPEEVRTP 517  
QY 560 SVCGFANPLPTESLDCFC-PHLQNSTOLEQVNOMLNLTQBEITATVKNVLPFGSPRYLOK 618  
Db 518 NYFGIMYLOSDDLGCCTCDDKVEPKKLDLNLKRLHTK----GSTERHLLYGRPAVLYR 573  
QY 619 NVDHCLLYHREYVSGFGKAMRMPMWSYTVPLGDTSELPPTVPDCLRADYRVPPSSSQK 678  
Db 574 -TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRSVSPFSQN 632  
QY 679 CSFYLDKNITHGFLYPPASNETSDOYDALITSNLVPMYEEFERKMDYFHSVLLIKHAT 738  
Db 633 CLAYKNDKOMSGLFFPYLSSPEAKYDAFLVNNVWYPAFRVWNYFORVLVKKYAS 692  
QY 739 ERNGNVVSGPIFDYNDYGHEDAPDEITKHLANTDVPITPHYFVVLTSCKNKSHTPENC 798  
Db 693 ERNGNVVSGPIFDYNDYGLHDTEDKIKQYVEGSSIPVPTHYSIITSCLDFTQPADKCD 752  
QY 799 GWLDVLPFIIPHRPTNVESCPGKPEALWVERFTAHTARVDRDVELLTGLDFYODKYQPV 858  
Db 753 GPLSVSSFILHRPDNEESCNSDESKWVELMKQHTARVDRDIEHLTSLDFFKTSRSY 812  
QY 859 SEILQLKTYLPTFTTI 875  
Db 813 PEILTLKTYLHTYESEI 829

Search completed: July 6, 2004, 13:21:04  
Job time : 27 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:19:53 ; Search time 56 Seconds  
(without alignments)  
4863.810 Million cell updates/sec

Title: US-10-005-480A-743  
Perfect score: 4804  
Sequence: 1 MESTLTATEQPVKNTLKK.....QPVSEIOLKYLPTFETI 875

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4804	100.0	875	14	US-10-062-109A-743
2	4804	100.0	875	14	US-10-062-109A-743
3	4804	100.0	875	14	US-10-062-109A-743
4	4804	100.0	875	14	US-10-005-480A-743
5	4804	100.0	875	14	US-10-005-480A-743
6	4804	100.0	875	14	US-10-005-480A-751
7	4804	100.0	875	15	US-10-291-241-3
8	4804	100.0	875	15	US-10-291-241-11
9	4804	100.0	875	15	US-10-291-241-16
10	4804	100.0	875	15	US-10-291-241-21
11	4804	100.0	875	15	US-10-291-241-22
12	4804	100.0	875	15	US-10-291-241-81
13	4804	100.0	875	15	US-10-291-241-94
14	4804	100.0	875	15	US-10-291-241-95
15	4804	100.0	875	15	US-10-291-241-96

16	4804	100.0	875	15	US-10-291-241-103
17	4801	99.9	875	14	US-10-062-109A-745
18	4801	99.9	875	14	US-10-062-109A-748
19	4801	99.9	875	14	US-10-005-480A-745
20	4801	99.9	875	14	US-10-005-480A-748
21	4801	99.9	875	15	US-10-291-241-5
22	4801	99.9	875	15	US-10-291-241-17
23	4798	99.9	875	15	US-10-291-241-9
24	4798	99.9	875	15	US-10-291-241-19
25	4797	99.9	875	15	US-10-291-241-7
26	4797	99.9	875	15	US-10-291-241-18
27	4796	99.8	875	15	US-10-291-241-13
28	4704	97.9	871	14	US-10-062-109A-765
29	4704	97.9	871	14	US-10-005-480A-765
30	4642	96.6	841	15	US-10-291-241-15
31	4642	96.6	841	15	US-10-291-241-101
32	4634	96.5	841	15	US-10-291-241-20
33	4634	96.5	841	15	US-10-291-241-100
34	4634	96.5	841	15	US-10-291-241-102
35	2498.5	52.0	873	9	US-09-997-701-6
36	2265	47.1	945	12	US-10-296-115-1102
37	2120	44.1	862	12	US-10-250-342-2
38	2108	43.9	862	12	US-10-307-817-603
39	2094.5	43.6	863	12	US-10-307-817-278
40	2094.5	43.6	863	12	US-10-307-817-604
41	2094.5	43.6	863	12	US-10-307-817-605
42	2094.5	43.6	863	12	US-10-307-817-610
43	2091.5	43.5	870	12	US-10-307-817-286
44	2091.5	43.5	871	12	US-10-307-817-276
45	2084	43.4	384	14	US-10-062-109A-749

ALIGNMENTS

RESULT 1

US-10-062-109A-743  
Sequence 743, Application US/10062109A  
Publication No. US20030165505A1  
GENERAL INFORMATION:  
APPLICANT: Ageneys  
APPLICANT: Chaltica-Bid, Pia M.  
APPLICANT: Raitano, Arthur B.  
APPLICANT: Paris, Mary  
APPLICANT: Hubert, Rene S.  
APPLICANT: Morrison, Karen Jane Meyrick  
APPLICANT: Jakobovits, Ava  
TITLE OF INVENTION: Nucleic Acid and Corresponding Protein  
TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of  
TITLE OF INVENTION: Cancer  
FILE REFERENCE: 51158-20062.01  
CURRENT APPLICATION NUMBER: US/10/062,109A  
CURRENT FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: US 10/005,480  
PRIOR FILING DATE: 2001-11-07  
NUMBER OF SEQ ID NOS: 765  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 743  
LENGTH: 875  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-062-109A-743

Query Match 100.0%; Score 4804; DB 14; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MESTLTATEQPVKNTLKKYKIACTVLLALLVIMSLGGLGRLKLEKQSCRRKCFD	60
Db	1	MESTLTATEQPVKNTLKKYKIACTVLLALLVIMSLGGLGRLKLEKQSCRRKCFD	60
Qy	61	ASPRGLENCRDVACKDRGDCWDFEDTQVESTRIWMCNFKFCGETRLRASI	120

```

Db 61 ASFRGLENCRCVACKDRCDCWDFEDTCVETRIWMCNKFRCGETRLEASLCSDDCL 120
QY 121 OKKCCADYKVCQGETSWLEENCDTAQOQSCPEGFOLPPVILFMDGFRAEYLYTWDTL 180
Db 121 OKKCCADYKVCQGETSWLEENCDTAQOQSCPEGFOLPPVILFMDGFRAEYLYTWDTL 180
QY 181 MPNINKLTCGHSKYRAMPYPTKTFNHYTIVTGLYPESHGIIDNNMYDNLNKNFSL 240
Db 181 MPNINKLTCGHSKYRAMPYPTKTFNHYTIVTGLYPESHGIIDNNMYDNLNKNFSL 240
QY 241 SKEQNNPAWHGQPMWLTAMTQGLKAATYFWPGSEVAINGSFSPSIYMPYNGSVPFPERIS 300
Db 241 SKEQNNPAWHGQPMWLTAMTQGLKAATYFWPGSEVAINGSFSPSIYMPYNGSVPFPERIS 300
QY 301 TLLKWLDPKAEPRFYMTYFEEDSSGHAGGPPVSARVICALQVDFHAFGLMEGLKQRN 360
Db 301 TLLKWLDPKAEPRFYMTYFEEDSSGHAGGPPVSARVICALQVDFHAFGLMEGLKQRN 360
QY 361 LHNCVNIILLADHGMDOQTCNKMEYMTDYPFRINFFYMEGPAAPRAHNI PHDFFSFS 420
Db 361 LHNCVNIILLADHGMDOQTCNKMEYMTDYPFRINFFYMEGPAAPRAHNI PHDFFSFS 420
QY 421 BEI VRNLSCKRPDOHFKPYLTDPDL PKLHYAKNVRIDKVLH FVDQQLAVRSKSNNTCCG 480
Db 421 BEI VRNLSCKRPDOHFKPYLTDPDL PKLHYAKNVRIDKVLH FVDQQLAVRSKSNNTCCG 480
QY 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540
Db 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540
QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVQMLNLTQBEI 600
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVQMLNLTQBEI 600
QY 601 TATVKNLPGRRVLQKNVDHCLLYHREYVSGFGKAMPMPWSSYTVPLQDTSPLPPT 660
Db 601 TATVKNLPGRRVLQKNVDHCLLYHREYVSGFGKAMPMPWSSYTVPLQDTSPLPPT 660
QY 661 VPCLRADVRVPSESKFSYFLADKNITHGFLYPPASNRTSDSQYDALITSNLVPWYEE 720
Db 661 VPCLRADVRVPSESKFSYFLADKNITHGFLYPPASNRTSDSQYDALITSNLVPWYEE 720
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDFADEITKHLANTDVP1PTHY 780
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDFADEITKHLANTDVP1PTHY 780
QY 781 FVVLTSCKNKSHTPENCPCWLDVLPPII PHRTNVSCEGKPEALWVEERTAHIAVR 840
Db 781 FVVLTSCKNKSHTPENCPCWLDVLPPII PHRTNVSCEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQPVSEIILQLKTYLPTFTTI 875
Db 841 DVELLTGLDFYQDKVQPVSEIILQLKTYLPTFTTI 875

```

## RESULT 2

```

US-10-062-109A-747
; Sequence 747, Application US/10062109A
; Publication No. US20030165505A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Chailita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen
; APPLICANT: Jane Meyrick
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P210B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20062.01
; CURRENT APPLICATION NUMBER: US/10/062,109A
; CURRENT FILING DATE: 2002-01-31

```

```

; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 747
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-109A-747

```

```

Query Match      100.0%; Score 4804; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MESTLTLATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLKLEKQSGCRKCCFD 60
Db 1 MESTLTLATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLKLEKQSGCRKCCFD 60
QY 61 ASFRGLENCRCVACKDRCDCWDFEDTCVETRIWMCNKFRCGETRLEASLCSDDCL 120
Db 61 ASFRGLENCRCVACKDRCDCWDFEDTCVETRIWMCNKFRCGETRLEASLCSDDCL 120
QY 121 OKKCCADYKVCQGETSWLEENCDTAQOQSCPEGFOLPPVILFMDGFRAEYLYTWDTL 180
Db 121 OKKCCADYKVCQGETSWLEENCDTAQOQSCPEGFOLPPVILFMDGFRAEYLYTWDTL 180
QY 181 MPNINKLTCGHSKYRAMPYPTKTFNHYTIVTGLYPESHGIIDNNMYDNLNKNFSL 240
Db 181 MPNINKLTCGHSKYRAMPYPTKTFNHYTIVTGLYPESHGIIDNNMYDNLNKNFSL 240
QY 241 SKEQNNPAWHGQPMWLTAMTQGLKAATYFWPGSEVAINGSFSPSIYMPYNGSVPFPERIS 300
Db 241 SKEQNNPAWHGQPMWLTAMTQGLKAATYFWPGSEVAINGSFSPSIYMPYNGSVPFPERIS 300
QY 301 TLLKWLDPKAEPRFYMTYFEEDSSGHAGGPPVSARVICALQVDFHAFGLMEGLKQRN 360
Db 301 TLLKWLDPKAEPRFYMTYFEEDSSGHAGGPPVSARVICALQVDFHAFGLMEGLKQRN 360
QY 361 LHNCVNIILLADHGMDOQTCNKMEYMTDYPFRINFFYMEGPAAPRAHNI PHDFFSFS 420
Db 361 LHNCVNIILLADHGMDOQTCNKMEYMTDYPFRINFFYMEGPAAPRAHNI PHDFFSFS 420
QY 421 BEI VRNLSCKRPDOHFKPYLTDPDL PKLHYAKNVRIDKVLH FVDQQLAVRSKSNNTCCG 480
Db 421 BEI VRNLSCKRPDOHFKPYLTDPDL PKLHYAKNVRIDKVLH FVDQQLAVRSKSNNTCCG 480
QY 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540
Db 481 GNHGYNNEFRSMEAI FLAHGPFSEKTEVEPFENIEVYNLMCDLLRIQAPNNGTHGSLN 540
QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVQMLNLTQBEI 600
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVQMLNLTQBEI 600
QY 601 TATVKNLPGRRVLQKNVDHCLLYHREYVSGFGKAMPMPWSSYTVPLQDTSPLPPT 660
Db 601 TATVKNLPGRRVLQKNVDHCLLYHREYVSGFGKAMPMPWSSYTVPLQDTSPLPPT 660
QY 661 VPCLRADVRVPSESKFSYFLADKNITHGFLYPPASNRTSDSQYDALITSNLVPWYEE 720
Db 661 VPCLRADVRVPSESKFSYFLADKNITHGFLYPPASNRTSDSQYDALITSNLVPWYEE 720
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDFADEITKHLANTDVP1PTHY 780
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDFADEITKHLANTDVP1PTHY 780
QY 781 FVVLTSCKNKSHTPENCPCWLDVLPPII PHRTNVSCEGKPEALWVEERTAHIAVR 840
Db 781 FVVLTSCKNKSHTPENCPCWLDVLPPII PHRTNVSCEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQPVSEIILQLKTYLPTFTTI 875
Db 841 DVELLTGLDFYQDKVQPVSEIILQLKTYLPTFTTI 875

```

## RESULT 3

US-10-062-109A-751  
; Sequence 751, Application US/10062109A  
; Publication No. US20030165505A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein  
; TITLE OF INVENTION: Entitled 161P2F108 Useful in Treatment and Detection of  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 51158-20062.01  
; CURRENT APPLICATION NUMBER: US/10/062.109A  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 10/005,480  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 765  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 751  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-062-109A-751

Query Match 100.0%; Score 4804; DB 14; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESTLTATEOPVKKNTLKKYKACIVLLALLVIMSLGLGLGLRLKLEKQSGCRKKCFD 60  
Db 1 MESTLTATEOPVKKNTLKKYKACIVLLALLVIMSLGLGLGLRLKLEKQSGCRKKCFD 60  
QY 61 ASFRGLENCRCVACKRGDCDCCWDFEDTCVESTRIWMCNKFRCGEIRLEASLCSDDCL 120  
Db 61 ASFRGLENCRCVACKRGDCDCCWDFEDTCVESTRIWMCNKFRCGEIRLEASLCSDDCL 120  
QY 121 QKCCADYKSVCOGETSWLEENCDTAQSCQCEGFDLPVILFSDMGFAEYLYTWDTL 180  
Db 121 QKCCADYKSVCOGETSWLEENCDTAQSCQCEGFDLPVILFSDMGFAEYLYTWDTL 180  
QY 181 MENINKLKTGCIHSHKYMEAYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240  
Db 181 MENINKLKTGCIHSHKYMEAYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240  
QY 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVFPERIS 300  
Db 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVFPERIS 300  
QY 301 TLLKWLDPKAEPRFRFTMYFEEDSSGHAGGPVSARVIALQVVDHAFGLMEGLKQRN 360  
Db 301 TLLKWLDPKAEPRFRFTMYFEEDSSGHAGGPVSARVIALQVVDHAFGLMEGLKQRN 360  
QY 361 LNCVNIILLADHGMQTYCNKMEYMTDYPRIINFFYWGAPRIAHNIPHDFFSFS 420  
Db 361 LNCVNIILLADHGMQTYCNKMEYMTDYPRIINFFYWGAPRIAHNIPHDFFSFS 420  
QY 421 EIVENLSCKPDQHFKEPLTPDLPKRLHYAKNVRIDKVLFDVQOWLAVRKSNNTNCG 480  
Db 421 EIVENLSCKPDQHFKEPLTPDLPKRLHYAKNVRIDKVLFDVQOWLAVRKSNNTNCG 480  
QY 481 GNHGYNFRSMEALFLAHGSPFKEKTEVEPFENIEVNLKCDLLRTQAPNNGTHGSLN 540  
Db 481 GNHGYNFRSMEALFLAHGSPFKEKTEVEPFENIEVNLKCDLLRTQAPNNGTHGSLN 540  
QY 541 HLLKVPFEPHABEVSFSGFANPLPTESLDPCFHLQNSQLQVQNMNLNLTQEEI 600  
Db 541 HLLKVPFEPHABEVSFSGFANPLPTESLDPCFHLQNSQLQVQNMNLNLTQEEI 600

QY 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVQLGDTSLPPT 660  
Db 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVQLGDTSLPPT 660  
QY 661 VPDCLRADVRVPSESQKCSFYLDKNIITGFLYPPASNRSTSQYDALITSNLVPMYEE 720  
Db 661 VPDCLRADVRVPSESQKCSFYLDKNIITGFLYPPASNRSTSQYDALITSNLVPMYEE 720  
QY 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPITPHY 780  
Db 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPITPHY 780  
QY 781 FVLTSCNKSHTPENCGLWDLVLPFIIPHRPTNVWSCPEKPEALWVERFTHARVR 840  
Db 781 FVLTSCNKSHTPENCGLWDLVLPFIIPHRPTNVWSCPEKPEALWVERFTHARVR 840  
QY 841 DVELLTGLDFYQDKVQVSVSEILQLKTLPTFETI 875  
Db 841 DVELLTGLDFYQDKVQVSVSEILQLKTLPTFETI 875

## RESULT 4

US-10-005-480A-743  
; Sequence 743, Application US/10005480A  
; Publication No. US20030191073A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein  
; TITLE OF INVENTION: Entitled 161P2F108 Useful in Treatment and Detection of  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 51158-20062.00  
; CURRENT APPLICATION NUMBER: US/10/005,480A  
; CURRENT FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 765  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 743  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-005-480A-743

Query Match 100.0%; Score 4804; DB 14; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESTLTATEOPVKKNTLKKYKACIVLLALLVIMSLGLGLGLRLKLEKQSGCRKKCFD 60  
Db 1 MESTLTATEOPVKKNTLKKYKACIVLLALLVIMSLGLGLGLRLKLEKQSGCRKKCFD 60  
QY 61 ASFRGLENCRCVACKRGDCDCCWDFEDTCVESTRIWMCNKFRCGEIRLEASLCSDDCL 120  
Db 61 ASFRGLENCRCVACKRGDCDCCWDFEDTCVESTRIWMCNKFRCGEIRLEASLCSDDCL 120  
QY 121 QKCCADYKSVCOGETSWLEENCDTAQSCQCEGFDLPVILFSDMGFAEYLYTWDTL 180  
Db 121 QKCCADYKSVCOGETSWLEENCDTAQSCQCEGFDLPVILFSDMGFAEYLYTWDTL 180  
QY 181 MENINKLKTGCIHSHKYMEAYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240  
Db 181 MENINKLKTGCIHSHKYMEAYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDVNLKNFSL 240  
QY 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVFPERIS 300  
Db 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVFPERIS 300  
QY 301 TLLKWLDPKAEPRFRFTMYFEEDSSGHAGGPVSARVIALQVVDHAFGLMEGLKQRN 360

```

Db 301 TLLKWLDDLPKAEPRFYTYMEEDSSGHAGGPVSARVVKALQVVDHAFQMLMEGLKQRN 360
Qy 361 LHNCVNIILLADHGMDOQTYCNKMEYMTDYPRINFFYMYEGPAPRIAHNI PHDFFSFNS 420
Db 361 LHNCVNIILLADHGMDOQTYCNKMEYMTDYPRINFFYMYEGPAPRIAHNI PHDFFSFNS 420
Qy 421 BEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVDOQLAVRSKNTNCGG 480
Db 421 BEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVDOQLAVRSKNTNCGG 480
Qy 481 GNGHYNNEFRSMEALFLAHGSPFKTEVEPFENIEVYNLMCDLLRIQAPANNTHGSLN 540
Db 481 GNGHYNNEFRSMEALFLAHGSPFKTEVEPFENIEVYNLMCDLLRIQAPANNTHGSLN 540
Qy 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTQLEQVNMNLNLTQBEI 600
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTQLEQVNMNLNLTQBEI 600
Qy 601 TATVKVNLPPGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSSYTVPOLGDTSLPPT 660
Db 601 TATVKVNLPPGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSSYTVPOLGDTSLPPT 660
Qy 661 VPDCLRADVRVPPESESQKCSFYLDKNI THGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
Db 661 VPDCLRADVRVPPESESQKCSFYLDKNI THGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
Qy 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDADPEITKHLANTDVPITPHY 780
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDADPEITKHLANTDVPITPHY 780
Qy 781 FVLTSCNKSHTPENCPCWGLDVLFPFIIPHRPTNVESCPGKPEALWVEERFTHARVR 840
Db 781 FVLTSCNKSHTPENCPCWGLDVLFPFIIPHRPTNVESCPGKPEALWVEERFTHARVR 840
Qy 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875
Db 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875

RESULT 5
US-10-005-480A-747
; Sequence 747, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20062.00
; CURRENT APPLICATION NUMBER: US/10/005,480A
; CURRENT FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 747
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-005-480A-747

Query Match 100.0%; Score 4804; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESTLTATEQPVKNTLKKYKACIACIVLLALLVMSIGLGLGLKLEKQSCRRKCFD 60
Db 1 MESTLTATEQPVKNTLKKYKACIACIVLLALLVMSIGLGLGLKLEKQSCRRKCFD 60

```

```

Qy 61 ASFRGLENCRCVACKDRGCCWDFEDTCVESTRIIMCNKFRCGETRLBASICSDDCL 120
Db 61 ASFRGLENCRCVACKDRGCCWDFEDTCVESTRIIMCNKFRCGETRLBASICSDDCL 120
Qy 121 QKXCCADYKSVCGGETSWLEENCDAQOSQCEGFDLPVILFSDMGFRAEVLVYTWDTL 180
Db 121 QKXCCADYKSVCGGETSWLEENCDAQOSQCEGFDLPVILFSDMGFRAEVLVYTWDTL 180
Qy 181 MPNINKLKTGCIHSHKYNRAMYPKTPNNHYTIVTGLYPESHGIIIDNNMYDVNLMNFSLS 240
Db 181 MPNINKLKTGCIHSHKYNRAMYPKTPNNHYTIVTGLYPESHGIIIDNNMYDVNLMNFSLS 240
Qy 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFPWQSEVAINGSPESIYMPYNGSVPEERIS 300
Db 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFPWQSEVAINGSPESIYMPYNGSVPEERIS 300
Qy 301 TLLKWLDDLPKAEPRFYTYMEEDSSGHAGGPVSARVVKALQVVDHAFQMLMEGLKQRN 360
Db 301 TLLKWLDDLPKAEPRFYTYMEEDSSGHAGGPVSARVVKALQVVDHAFQMLMEGLKQRN 360
Qy 361 LHNCVNIILLADHGMDOQTYCNKMEYMTDYPRINFFYMYEGPAPRIAHNI PHDFFSFNS 420
Db 361 LHNCVNIILLADHGMDOQTYCNKMEYMTDYPRINFFYMYEGPAPRIAHNI PHDFFSFNS 420
Qy 421 BEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVDOQLAVRSKNTNCGG 480
Db 421 BEIVRNLSCKRPDQHFKPYLTPDLPKRLHYAKNVRIDKVLHFDVDOQLAVRSKNTNCGG 480
Qy 481 GNGHYNNEFRSMEALFLAHGSPFKTEVEPFENIEVYNLMCDLLRIQAPANNTHGSLN 540
Db 481 GNGHYNNEFRSMEALFLAHGSPFKTEVEPFENIEVYNLMCDLLRIQAPANNTHGSLN 540
Qy 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTQLEQVNMNLNLTQBEI 600
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCPHLQNSTQLEQVNMNLNLTQBEI 600
Qy 601 TATVKVNLPPGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSSYTVPOLGDTSLPPT 660
Db 601 TATVKVNLPPGRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSSYTVPOLGDTSLPPT 660
Qy 661 VPDCLRADVRVPPESESQKCSFYLDKNI THGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
Db 661 VPDCLRADVRVPPESESQKCSFYLDKNI THGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
Qy 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDADPEITKHLANTDVPITPHY 780
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHDADPEITKHLANTDVPITPHY 780
Qy 781 FVLTSCNKSHTPENCPCWGLDVLFPFIIPHRPTNVESCPGKPEALWVEERFTHARVR 840
Db 781 FVLTSCNKSHTPENCPCWGLDVLFPFIIPHRPTNVESCPGKPEALWVEERFTHARVR 840
Qy 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875
Db 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875

RESULT 6
US-10-005-480A-751
; Sequence 751, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer

```

```
; FILE REFERENCE: 51158-20062.00
; CURRENT APPLICATION NUMBER: US/10/005,480A
; CURRENT FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 751
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-005-480A-751

Query Match      100.0%; Score 4804; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
Db 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
QY 61 ASFRGLENCRCVACKDRGCCWDFEDTCVSTRIMWKNKFCGTRLEASLCSDDCL 120
Db 61 ASFRGLENCRCVACKDRGCCWDFEDTCVSTRIMWKNKFCGTRLEASLCSDDCL 120
QY 121 QKDCADYKSVCOGETSWLENCDDTAQSCQCEGFDLPVILFSDMGFRAEYLTYWDTL 180
Db 121 QKDCADYKSVCOGETSWLENCDDTAQSCQCEGFDLPVILFSDMGFRAEYLTYWDTL 180
QY 181 MENINKLTCGTHSKYWEAMYPKTPTFNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
Db 181 MENINKLTCGTHSKYWEAMYPKTPTFNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
QY 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300
Db 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300
QY 301 TLLKWLDPKAEPRFYTMWFEEDPSGSHAGGVPVSARVVKALQVVDHAFGLMEGLKQRN 360
Db 301 TLLKWLDPKAEPRFYTMWFEEDPSGSHAGGVPVSARVVKALQVVDHAFGLMEGLKQRN 360
QY 361 LHNCVNIILLADHGMQDTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFSFNS 420
Db 361 LHNCVNIILLADHGMQDTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFSFNS 420
QY 421 EEIVRLNSCRKPDQHFKPYLTPDLKRLHYAKNVRIDKVLHFVDQQWLAVRSKSNTCGG 480
Db 421 EEIVRLNSCRKPDQHFKPYLTPDLKRLHYAKNVRIDKVLHFVDQQWLAVRSKSNTCGG 480
QY 481 GNEGYNNEFRSMEAI FLAHCPSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540
Db 481 GNEGYNNEFRSMEAI FLAHCPSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540
QY 541 HLKVPPEPSHABEYSKFSVCGFANPLPTESLDCFCPLONSTQLEQVNMNLNLTQEEI 600
Db 541 HLKVPPEPSHABEYSKFSVCGFANPLPTESLDCFCPLONSTQLEQVNMNLNLTQEEI 600
QY 601 TATVKVNLPGRPVLOKQNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSLPPT 660
Db 601 TATVKVNLPGRPVLOKQNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSLPPT 660
QY 661 VPDCLRADVPPSESKCSFYADKNITHGFLYPPASNTSDSYDALITSNLVPMWYEE 720
Db 661 VPDCLRADVPPSESKCSFYADKNITHGFLYPPASNTSDSYDALITSNLVPMWYEE 720
QY 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNYDGHFADPDEITKHLANTDVPITPHY 780
Db 721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNYDGHFADPDEITKHLANTDVPITPHY 780
QY 781 FVLTSCNKSHTPENCGLDVLPTIIPHRPTNVESCPEGKPEALWVERFRTAHTARVR 840
Db 781 FVLTSCNKSHTPENCGLDVLPTIIPHRPTNVESCPEGKPEALWVERFRTAHTARVR 840
QY 841 DVELLTGLDFYQDKVPVSEILQLKTYLPTFETI 875
Db 841 DVELLTGLDFYQDKVPVSEILQLKTYLPTFETI 875
```

```
Db 841 DVELLTGLDFYQDKVPVSEILQLKTYLPTFETI 875

RESULT 7
US-10-291-241-3
; Sequence 3, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Weyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F:0B USEFUL IN TREATMENT AND DETECTION OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-3
```

```
Query Match      100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
Db 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60
QY 61 ASFRGLENCRCVACKDRGCCWDFEDTCVSTRIMWKNKFCGTRLEASLCSDDCL 120
Db 61 ASFRGLENCRCVACKDRGCCWDFEDTCVSTRIMWKNKFCGTRLEASLCSDDCL 120
QY 121 QKDCADYKSVCOGETSWLENCDDTAQSCQCEGFDLPVILFSDMGFRAEYLTYWDTL 180
Db 121 QKDCADYKSVCOGETSWLENCDDTAQSCQCEGFDLPVILFSDMGFRAEYLTYWDTL 180
QY 181 MENINKLTCGTHSKYWEAMYPKTPTFNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
Db 181 MENINKLTCGTHSKYWEAMYPKTPTFNHYTIVTGLYPESHGIIIDNNMYDVLNKNFSL 240
QY 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300
Db 241 SKEQNPAAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPEERIS 300
QY 301 TLLKWLDPKAEPRFYTMWFEEDPSGSHAGGVPVSARVVKALQVVDHAFGLMEGLKQRN 360
Db 301 TLLKWLDPKAEPRFYTMWFEEDPSGSHAGGVPVSARVVKALQVVDHAFGLMEGLKQRN 360
QY 361 LHNCVNIILLADHGMQDTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFSFNS 420
Db 361 LHNCVNIILLADHGMQDTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFSFNS 420
QY 421 EEIVRLNSCRKPDQHFKPYLTPDLKRLHYAKNVRIDKVLHFVDQQWLAVRSKSNTCGG 480
Db 421 EEIVRLNSCRKPDQHFKPYLTPDLKRLHYAKNVRIDKVLHFVDQQWLAVRSKSNTCGG 480
```



```
QY 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPAPNNGTHGSLN 540
DB 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPAPNNGTHGSLN 540
QY 541 HLLKVPFYSHAEVSKFVCGFANPLPTESLDCFCPLHONSTQLEOVNOMLNLTOEEI 600
DB 541 HLLKVPFYSHAEVSKFVCGFANPLPTESLDCFCPLHONSTQLEOVNOMLNLTOEEI 600
QY 601 TATVKVNLPRGRPRVLQKNDHCLLYHREYVSGFGKAMPMMSSYTVPLQDTSPLPPT 660
DB 601 TATVKVNLPRGRPRVLQKNDHCLLYHREYVSGFGKAMPMMSSYTVPLQDTSPLPPT 660
QY 661 VPCLRADVRVPSESKCSFYLDKXNITHGFLYPPASNTSDQDALITSNLVPMEYE 720
DB 661 VPCLRADVRVPSESKCSFYLDKXNITHGFLYPPASNTSDQDALITSNLVPMEYE 720
QY 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPPIPTHY 780
DB 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPPIPTHY 780
QY 781 FVLTSCNKSHTPENCPCGWLVDLPIIHRPTNVSCPEGKPEALWVEERTAHIAVR 840
DB 781 FVLTSCNKSHTPENCPCGWLVDLPIIHRPTNVSCPEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQVPSVSEILQKTYLPTFETI 875
DB 841 DVELLTGLDFYQDKVQVPSVSEILQKTYLPTFETI 875

RESULT 8
US-10-291-241-11
; Sequence 11, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; PRIOR FILING DATE: 2003-03-15
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-11

Query Match 100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPKVKNLTKYKACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
DB 1 MESTLTATEQPKVKNLTKYKACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
QY 61 ASFRGLENCRCDVACKRGDCWDFEDTCVSTRIMCNKFRCGTRELASLCSGSDCL 120
```

```
DB 61 ASFRGLENCRCDVACKRGDCWDFEDTCVSTRIMCNKFRCGTRELASLCSGSDCL 120
QY 121 QKKDCADYKVCQGETSWLEBENDTAQOSCPGEGFDLPPVILFSMDGFAEYLYTWDTL 180
DB 121 QKKDCADYKVCQGETSWLEBENDTAQOSCPGEGFDLPPVILFSMDGFAEYLYTWDTL 180
QY 181 MPNINKLTCGIGHSKYWRAMYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLKNFSL 240
DB 181 MPNINKLTCGIGHSKYWRAMYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLKNFSL 240
QY 241 SKEQNPAAWHGQPMWLTAMYQG:KAATYFWPGSEVAINGSPSIYMPYNGSVPFERIS 300
DB 241 SKEQNPAAWHGQPMWLTAMYQG:KAATYFWPGSEVAINGSPSIYMPYNGSVPFERIS 300
QY 301 TLLKWLDPKAEPRFYMTYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQRN 360
DB 301 TLLKWLDPKAEPRFYMTYFEEPDSSGHAGGPVSARVIKALQVVDHAFGLMEGLKQRN 360
QY 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFYMYEGPAPRIAHNI PHDFFSFS 420
DB 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFYMYEGPAPRIAHNI PHDFFSFS 420
QY 421 BEIVRNLSCKRPDQHFPRYLTPOLPKELHYAKNVRIDKVLHFDVDDQWLAVRSKNTNCGG 480
DB 421 BEIVRNLSCKRPDQHFPRYLTPOLPKELHYAKNVRIDKVLHFDVDDQWLAVRSKNTNCGG 480
QY 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPAPNNGTHGSLN 540
DB 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVEPFENIEVYNLMCDLLRIQAPAPNNGTHGSLN 540
QY 541 HLLKVPFYSHAEVSKFVCGFANPLPTESLDCFCPLHONSTQLEOVNOMLNLTOEEI 600
DB 541 HLLKVPFYSHAEVSKFVCGFANPLPTESLDCFCPLHONSTQLEOVNOMLNLTOEEI 600
QY 601 TATVKVNLPRGRPRVLQKNDHCLLYHREYVSGFGKAMPMMSSYTVPLQDTSPLPPT 660
DB 601 TATVKVNLPRGRPRVLQKNDHCLLYHREYVSGFGKAMPMMSSYTVPLQDTSPLPPT 660
QY 661 VPCLRADVRVPSESKCSFYLDKXNITHGFLYPPASNTSDQDALITSNLVPMEYE 720
DB 661 VPCLRADVRVPSESKCSFYLDKXNITHGFLYPPASNTSDQDALITSNLVPMEYE 720
QY 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPPIPTHY 780
DB 721 FRKWDYFHSVLLIKHATERGVNVVSGPIFDYNDGHFADPDEITKHLANTDVPPIPTHY 780
QY 781 FVLTSCNKSHTPENCPCGWLVDLPIIHRPTNVSCPEGKPEALWVEERTAHIAVR 840
DB 781 FVLTSCNKSHTPENCPCGWLVDLPIIHRPTNVSCPEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQVPSVSEILQKTYLPTFETI 875
DB 841 DVELLTGLDFYQDKVQVPSVSEILQKTYLPTFETI 875

RESULT 9
US-10-291-241-16
; Sequence 16, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; FILE REFERENCE: 51158-20062.20
```

```
; CURRENT APPLICATION NUMBER: US/10/291,241
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-16

Query Match      100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
DB 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
QY 61 ASFRGLENCRCVACKDRCDCWDFEDTCTVSTRIMWKNKFCGTRLEASLCSDDCL 120
DB 61 ASFRGLENCRCVACKDRCDCWDFEDTCTVSTRIMWKNKFCGTRLEASLCSDDCL 120
QY 121 QKQCCADYKSVCOGSETSWLENCDDTAQSQCPGDFLPPVILFSMDGFRAYLYTWDTL 180
DB 121 QKQCCADYKSVCOGSETSWLENCDDTAQSQCPGDFLPPVILFSMDGFRAYLYTWDTL 180
QY 181 MNINKLTCGTHSKYMRAMYPKTPFNHYITVTGLYPESHGIIDNNYDVNLKNFSL 240
DB 181 MNINKLTCGTHSKYMRAMYPKTPFNHYITVTGLYPESHGIIDNNYDVNLKNFSL 240
QY 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPEERIS 300
DB 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPEERIS 300
QY 301 TLLKWLDPKAEPRFYTYMFEEDSSGHAGGVPVSARVICALQVVDHAFGLMEGLKQ 360
DB 301 TLLKWLDPKAEPRFYTYMFEEDSSGHAGGVPVSARVICALQVVDHAFGLMEGLKQ 360
QY 361 LHCNVNIIILLADHGMQTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFFS 420
DB 361 LHCNVNIIILLADHGMQTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFFS 420
QY 421 EEIVRLSCRKPDQHPKPYLTPDLPKRLHYAKNVRIDKVHLFVDOQWLAVRSKNTNCG 480
DB 421 EEIVRLSCRKPDQHPKPYLTPDLPKRLHYAKNVRIDKVHLFVDOQWLAVRSKNTNCG 480
QY 481 GNHGYNNEFSEMEALFLAHGSPFKEKTEVEPPENIEVNIIMCDLLRIQAPNNGTHGSLN 540
DB 481 GNHGYNNEFSEMEALFLAHGSPFKEKTEVEPPENIEVNIIMCDLLRIQAPNNGTHGSLN 540
QY 541 HLLKVPFYPESHABEVSFKSVCGFANPLPTESLDCFCPLHNSQLEQVNMNLTOEEI 600
DB 541 HLLKVPFYPESHABEVSFKSVCGFANPLPTESLDCFCPLHNSQLEQVNMNLTOEEI 600
QY 601 TATVKVNLPGRRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPLGDTSLPPT 660
DB 601 TATVKVNLPGRRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPLGDTSLPPT 660
QY 661 VPDCLRADVPPSESKOCSFYLLADKNI THGFLYPPASNETSDSQYDALITSNLVPMYEE 720
DB 661 VPDCLRADVPPSESKOCSFYLLADKNI THGFLYPPASNETSDSQYDALITSNLVPMYEE 720
QY 721 FRKMWDYFHSVLLIKHATERGVNVVSGPIFDYNYDGHFADPEITKHLANTDVPITPHY 780
DB 721 FRKMWDYFHSVLLIKHATERGVNVVSGPIFDYNYDGHFADPEITKHLANTDVPITPHY 780
QY 781 FVLTSCNKSHTPENCPGMLDVLPII PHRPTNVESCPGKPEALWVERFTAHIAVR 840
```

```
DB 781 FVLTSCNKSHTPENCPGMLDVLPII PHRPTNVESCPGKPEALWVERFTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQVPSVSEILOKTYLPTFETTI 875
DB 841 DVELLTGLDFYQDKVQVPSVSEILOKTYLPTFETTI 875

RESULT 10
US-10-291-241-21
; Sequence 21, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 1612F108 USEFUL IN TREATMENT AND DETECTION OF
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-21

Query Match      100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
DB 1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRLEKQSGCRKKCFD 60
QY 61 ASFRGLENCRCVACKDRCDCWDFEDTCTVSTRIMWKNKFCGTRLEASLCSDDCL 120
DB 61 ASFRGLENCRCVACKDRCDCWDFEDTCTVSTRIMWKNKFCGTRLEASLCSDDCL 120
QY 121 QKQCCADYKSVCOGSETSWLENCDDTAQSQCPGDFLPPVILFSMDGFRAYLYTWDTL 180
DB 121 QKQCCADYKSVCOGSETSWLENCDDTAQSQCPGDFLPPVILFSMDGFRAYLYTWDTL 180
QY 181 MNINKLTCGTHSKYMRAMYPKTPFNHYITVTGLYPESHGIIDNNYDVNLKNFSL 240
DB 181 MNINKLTCGTHSKYMRAMYPKTPFNHYITVTGLYPESHGIIDNNYDVNLKNFSL 240
QY 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPEERIS 300
DB 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGEVAINGSPFSIYMPYNGSVPEERIS 300
QY 301 TLLKWLDPKAEPRFYTYMFEEDSSGHAGGVPVSARVICALQVVDHAFGLMEGLKQ 360
DB 301 TLLKWLDPKAEPRFYTYMFEEDSSGHAGGVPVSARVICALQVVDHAFGLMEGLKQ 360
QY 361 LHCNVNIIILLADHGMQTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFFS 420
DB 361 LHCNVNIIILLADHGMQTYCNKMEYMTDYPFRINFFYMEGAPRIRAHNI PHDFFS 420
```

```
QY 421 EEI VRNLS CRKPDQHF KPYLT PDL PKELHYAKNVRIDKVHLFVDQQLAVRSKSNNTCCG 480
DB 421 EEI VRNLS CRKPDQHF KPYLT PDL PKELHYAKNVRIDKVHLFVDQQLAVRSKSNNTCCG 480
QY 481 GNEGYNNEFRSMEAI FLAHGSPSKEKTEVEPFENIEVYNLMCDLRIQAPANNNGTHGSLN 540
DB 481 GNEGYNNEFRSMEAI FLAHGSPSKEKTEVEPFENIEVYNLMCDLRIQAPANNNGTHGSLN 540
QY 541 HLLKVPFEPESHAEVSKFSVCGFANPLPTESIDCFPHLQNSTQLEQVNMQLNLTQEEI 600
DB 541 HLLKVPFEPESHAEVSKFSVCGFANPLPTESIDCFPHLQNSTQLEQVNMQLNLTQEEI 600
QY 601 TATVKVNLPRGRVRLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660
DB 601 TATVKVNLPRGRVRLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660
QY 661 VPCLRADVRVPSESKCSFYLDKNI THGFLYPPASNRTSDSDALITSNLVPWYEE 720
DB 661 VPCLRADVRVPSESKCSFYLDKNI THGFLYPPASNRTSDSDALITSNLVPWYEE 720
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780
DB 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780
QY 781 FVVLTSCKNKSHTPENCPCGWLVDLPII PHRPNTNVESECEGKPEALWVEERTAHIAVR 840
DB 781 FVVLTSCKNKSHTPENCPCGWLVDLPII PHRPNTNVESECEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFFETI 875
DB 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFFETI 875

RESULT 11
US-10-291-241-22
; Sequence 22, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; PRIOR FILING DATE: 2003-03-15
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-22

Query Match 100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTLTLATEQPVKNTLUKKYKACIVLALLVMSLGLGLGLKLEKQSGCKKCFD 60
DB 61 ASFRGLENCRCVACKRGDCNDPFCVETSTRIWNCNFKRCGETRLEASLCSGSDCL 120
QY 61 ASFRGLENCRCVACKRGDCNDPFCVETSTRIWNCNFKRCGETRLEASLCSGSDCL 120
DB 121 QKDCADYKSVCCQGETSWLEENCDTAQSCPCGFDLPVILFSDMGFAEVLTYTWDTL 180
QY 121 QKDCADYKSVCCQGETSWLEENCDTAQSCPCGFDLPVILFSDMGFAEVLTYTWDTL 180
DB 181 MENINKLTCGISHKYNRAMPYTKTFPNHYTIVTGLYPESHGIIIDNNYDNLNKNFSL 240
QY 181 MENINKLTCGISHKYNRAMPYTKTFPNHYTIVTGLYPESHGIIIDNNYDNLNKNFSL 240
DB 241 SKEQNPANWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPFEEIS 300
QY 241 SKEQNPANWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPSIYMPYNGSVPFEEIS 300
DB 301 TLLKWLDPKPAERPRFTYMTYFEEPDSSGHAGGPVSARVICALQVVDHAFGLMEGLKQ 360
QY 301 TLLKWLDPKPAERPRFTYMTYFEEPDSSGHAGGPVSARVICALQVVDHAFGLMEGLKQ 360
DB 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPFRINFFYMYEGPAPRIAHNI PHDFSFS 420
QY 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPFRINFFYMYEGPAPRIAHNI PHDFSFS 420
DB 421 EEI VRNLS CRKPDQHF KPYLT PDL PKELHYAKNVRIDKVHLFVDQQLAVRSKSNNTCCG 480
QY 421 EEI VRNLS CRKPDQHF KPYLT PDL PKELHYAKNVRIDKVHLFVDQQLAVRSKSNNTCCG 480
DB 481 GNEGYNNEFRSMEAI FLAHGSPSKEKTEVEPFENIEVYNLMCDLRIQAPANNNGTHGSLN 540
QY 481 GNEGYNNEFRSMEAI FLAHGSPSKEKTEVEPFENIEVYNLMCDLRIQAPANNNGTHGSLN 540
DB 541 HLLKVPFEPESHAEVSKFSVCGFANPLPTESIDCFPHLQNSTQLEQVNMQLNLTQEEI 600
QY 541 HLLKVPFEPESHAEVSKFSVCGFANPLPTESIDCFPHLQNSTQLEQVNMQLNLTQEEI 600
DB 601 TATVKVNLPRGRVRLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660
QY 601 TATVKVNLPRGRVRLQKNVDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660
DB 661 VPCLRADVRVPSESKCSFYLDKNI THGFLYPPASNRTSDSDALITSNLVPWYEE 720
QY 661 VPCLRADVRVPSESKCSFYLDKNI THGFLYPPASNRTSDSDALITSNLVPWYEE 720
DB 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDVPITPHY 780
DB 781 FVVLTSCKNKSHTPENCPCGWLVDLPII PHRPNTNVESECEGKPEALWVEERTAHIAVR 840
QY 781 FVVLTSCKNKSHTPENCPCGWLVDLPII PHRPNTNVESECEGKPEALWVEERTAHIAVR 840
QY 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFFETI 875
DB 841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFFETI 875

RESULT 12
US-10-291-241-81
; Sequence 81, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
```

```
; APPLICANT: Pia M. Challita-Bid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US 10/291,241
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-81

Query Match      100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGURKLEKQSGCRKKCFD 60
Db      1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGURKLEKQSGCRKKCFD 60

QY      61 ASFRGLENCRCDVACKDRGCCWDFEDTCVESTRIWNCNFKRCGETRLEASCSDDC 120
Db      61 ASFRGLENCRCDVACKDRGCCWDFEDTCVESTRIWNCNFKRCGETRLEASCSDDC 120

QY      121 QKDCADYKSVCOGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEYLYTWD 180
Db      121 QKDCADYKSVCOGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEYLYTWD 180

QY      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240
Db      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240

QY      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240
Db      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240

QY      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300
Db      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300

QY      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300
Db      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300

QY      301 TLLKWLDPKAEPRFYTWYFEBPDSGSHAGGPVSARVIKALQVVDHAFGLMEGLKORN 360
Db      301 TLLKWLDPKAEPRFYTWYFEBPDSGSHAGGPVSARVIKALQVVDHAFGLMEGLKORN 360

QY      361 LHCNVIIILLADRGMDQTYCNKVEYMTDYPRINFFYMEGAPAPRAHNIHPDPFSNS 420
Db      361 LHCNVIIILLADRGMDQTYCNKVEYMTDYPRINFFYMEGAPAPRAHNIHPDPFSNS 420

QY      421 EEIVRNLSCKRPQHPKPYLTDPDKRLHYAKNVRIDKVLHFDVDOQWLAVRSKSNTCGG 480
Db      421 EEIVRNLSCKRPQHPKPYLTDPDKRLHYAKNVRIDKVLHFDVDOQWLAVRSKSNTCGG 480

QY      481 GNGHYNNFESMAEIFAHLGSPSPKETEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540
Db      481 GNGHYNNFESMAEIFAHLGSPSPKETEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540

QY      541 HLLKVPYFESHAEVSKFSVCGFANPLPTESLDCCPHLQNSQTOLEQVNMNLNTOEBE 600
Db      541 HLLKVPYFESHAEVSKFSVCGFANPLPTESLDCCPHLQNSQTOLEQVNMNLNTOEBE 600

QY      601 TATVKVNLPGCRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660
Db      601 TATVKVNLPGCRPRVLQKNDHCLLYHREYVSGFGKAMRPMWSSYTVPOLGDTSPLPPT 660

QY      661 VPDCLRADVVPSESOKCSFYLDKNI THGELYPPASNETSDQYDALITSLNLPWYEE 720
Db      661 VPDCLRADVVPSESOKCSFYLDKNI THGELYPPASNETSDQYDALITSLNLPWYEE 720
```

```
721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHFDAPDEITKHLANTDVPPIPHY 780
721 FRKWDYFHSVLLIKHATERNGVNVSGPIFDYNDGHFDAPDEITKHLANTDVPPIPHY 780

QY      781 FVULTSCNKSHTPENCPCGMLDVLFFIIPHRPTNVESCPGKPEALWVERFTHAIAVR 840
Db      781 FVULTSCNKSHTPENCPCGMLDVLFFIIPHRPTNVESCPGKPEALWVERFTHAIAVR 840

QY      841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875
Db      841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETI 875

RESULT 13
US-10-291-241-94
; Sequence 94, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Bid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-94

Query Match      100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGURKLEKQSGCRKKCFD 60
Db      1 MESTLTATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGURKLEKQSGCRKKCFD 60

QY      61 ASFRGLENCRCDVACKDRGCCWDFEDTCVESTRIWNCNFKRCGETRLEASCSDDC 120
Db      61 ASFRGLENCRCDVACKDRGCCWDFEDTCVESTRIWNCNFKRCGETRLEASCSDDC 120

QY      121 QKDCADYKSVCOGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEYLYTWD 180
Db      121 QKDCADYKSVCOGETSWLEENCDTAQSQCPGFDLPVILFSDMGFRAEYLYTWD 180

QY      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240
Db      181 MPNINKLTCGIIHSHKYNRMAYPTKTFPNHYITVGLYPESHGIIIDNNMYDVNLNKNFSL 240

QY      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300
Db      241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVVPEERIS 300

QY      301 TLLKWLDPKAEPRFYTWYFEBPDSGSHAGGPVSARVIKALQVVDHAFGLMEGLKORN 360
Db      301 TLLKWLDPKAEPRFYTWYFEBPDSGSHAGGPVSARVIKALQVVDHAFGLMEGLKORN 360
```

Db 301 TLLKWLDPKAEPRFYMTYFEBPDSSGHAGGVPASRVKALQVVDHAFGLMEGLKQRN 360  
QY 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMEGPAAPRAHNI PHDFFSFS 420  
Db 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMEGPAAPRAHNI PHDFFSFS 420  
QY 421 EETVRNLSCKRPQOHFKPYLTDPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNNTCCG 480  
Db 421 EETVRNLSCKRPQOHFKPYLTDPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNNTCCG 480  
QY 481 GNHGYNNEFRSMEAI FLAHGFSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Db 481 GNHGYNNEFRSMEAI FLAHGFSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCFHLQNSTQLEQVNMNLNLTQEEI 600  
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCFHLQNSTQLEQVNMNLNLTQEEI 600  
QY 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSYTVPLQGDTSPLPPT 660  
Db 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSYTVPLQGDTSPLPPT 660  
QY 661 VPCLRADVRVPSESKFSYLADKNI THGFLYPPASNRTSDSQYDALITSNLVPMBE 720  
Db 661 VPCLRADVRVPSESKFSYLADKNI THGFLYPPASNRTSDSQYDALITSNLVPMBE 720  
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
QY 781 FVLTSCNKSHTPENCPCGWLVDLPFIIPHRPNVSCPEGKPEALWVERFTAHIARVR 840  
Db 781 FVLTSCNKSHTPENCPCGWLVDLPFIIPHRPNVSCPEGKPEALWVERFTAHIARVR 840  
QY 841 DVELLTGLDFYQKQVPVSEILQKYLPTFTTI 875  
Db 841 DVELLTGLDFYQKQVPVSEILQKYLPTFTTI 875

## RESULT 14

US-10-291-241-95  
; Sequence 95, Application US/10291241  
; Publication No. US20030206905A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Aya Jakobovits  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Mary Faris  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Wangmao Ge  
; APPLICANT: Karen Jane Meyrick Morrison  
; APPLICANT: Robert Kendall Morrison  
; APPLICANT: Pia M. Challita-Eid  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 51158-20062.20  
; CURRENT APPLICATION NUMBER: US/10/291,241  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US 10/005,480  
; PRIOR FILING DATE: 2001-11-07  
; PRIOR APPLICATION NUMBER: US 10/062,109  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; NUMBER OF SEQ ID NOS: 103  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 95  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-291-241-95

Query Match 100.0%; Score 4804; DB 15; Length 875;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESTLTLATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60  
Db 1 MESTLTLATEQPVKNTLKKYKIACIVLLALLVIMSLGLGLGRKLEKQGSCKKCFD 60  
QY 61 ASFRGLENCRCVACKRGDCWDFEDTCVESTIWMCNKFRGETLEASLCSDDCL 120  
Db 61 ASFRGLENCRCVACKRGDCWDFEDTCVESTIWMCNKFRGETLEASLCSDDCL 120  
QY 121 QKQCCADYKVCQGETSWLEENCDTAQOQCPGFDLPVILFSDMGFRAEYLYTWDTL 180  
Db 121 QKQCCADYKVCQGETSWLEENCDTAQOQCPGFDLPVILFSDMGFRAEYLYTWDTL 180  
QY 181 MPNINKLTCGIHSHKMYRAMYPTKTFPNHYTITGLYPESHGIIIDNNMYDVNLNKNFSL 240  
Db 181 MPNINKLTCGIHSHKMYRAMYPTKTFPNHYTITGLYPESHGIIIDNNMYDVNLNKNFSL 240  
QY 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFPGSEVAINGSFPSIYMPYNGSVPFPERIS 300  
Db 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFPGSEVAINGSFPSIYMPYNGSVPFPERIS 300  
QY 301 TLLKWLDPKAEPRFYMTYFEBPDSSGHAGGVPASRVKALQVVDHAFGLMEGLKQRN 360  
Db 301 TLLKWLDPKAEPRFYMTYFEBPDSSGHAGGVPASRVKALQVVDHAFGLMEGLKQRN 360  
QY 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMEGPAAPRAHNI PHDFFSFS 420  
Db 361 LHNCVNIILLADHGMDQTYCNKMEYMTDYPRINFYMEGPAAPRAHNI PHDFFSFS 420  
QY 421 EETVRNLSCKRPQOHFKPYLTDPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNNTCCG 480  
Db 421 EETVRNLSCKRPQOHFKPYLTDPKRLHYAKNVRIDKVLHFDVQOQLAVRSKSNNTCCG 480  
QY 481 GNHGYNNEFRSMEAI FLAHGFSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
Db 481 GNHGYNNEFRSMEAI FLAHGFSFKEKTEVEPPENIEVYNLMCDLLRIQAPNNGTHGSLN 540  
QY 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCFHLQNSTQLEQVNMNLNLTQEEI 600  
Db 541 HLLKVPFYPESHAEVSKFSVCGFANPLPTESLDCCFHLQNSTQLEQVNMNLNLTQEEI 600  
QY 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSYTVPLQGDTSPLPPT 660  
Db 601 TATVKVNLPRGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSYTVPLQGDTSPLPPT 660  
QY 661 VPCLRADVRVPSESKFSYLADKNI THGFLYPPASNRTSDSQYDALITSNLVPMBE 720  
Db 661 VPCLRADVRVPSESKFSYLADKNI THGFLYPPASNRTSDSQYDALITSNLVPMBE 720  
QY 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
Db 721 FRKWDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
QY 781 FVLTSCNKSHTPENCPCGWLVDLPFIIPHRPNVSCPEGKPEALWVERFTAHIARVR 840  
Db 781 FVLTSCNKSHTPENCPCGWLVDLPFIIPHRPNVSCPEGKPEALWVERFTAHIARVR 840  
QY 841 DVELLTGLDFYQKQVPVSEILQKYLPTFTTI 875  
Db 841 DVELLTGLDFYQKQVPVSEILQKYLPTFTTI 875

## RESULT 15

US-10-291-241-96  
; Sequence 96, Application US/10291241  
; Publication No. US20030206905A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Aya Jakobovits  
; APPLICANT: Arthur B. Raitano

APPLICANT: Mary Paris  
APPLICANT: Rene S. Hubert  
APPLICANT: Wangmao Ge  
APPLICANT: Karen Jane Meyrick Morrison  
APPLICANT: Robert Kendall Morrison  
APPLICANT: Pia M. Challita-Eid  
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF  
TITLE OF INVENTION: CANCER  
FILE REFERENCE: 51158-20062.20  
CURRENT APPLICATION NUMBER: US/10/291,241  
CURRENT FILING DATE: 2003-03-18  
PRIOR APPLICATION NUMBER: US 10/005,480  
PRIOR FILING DATE: 2001-11-07  
PRIOR APPLICATION NUMBER: US 10/062,109  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: US 60/282,739  
PRIOR FILING DATE: 2001-04-10  
NUMBER OF SEQ ID NOS: 103  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 96  
LENGTH: 875  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-291-241-96

Query Match 100.0%; Score 4804; DB 15; Length 875;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESTLTATBOPVKNKTKKYYKACIVLLVIMSLGLGLGRKLEKQSCRKCPD 60  
Db 1 MESTLTATBOPVKNKTKKYYKACIVLLVIMSLGLGLGRKLEKQSCRKCPD 60  
Qy 61 ASFPLENCRCDVACKDRCDCDFEDTCVETRIWNCNFRGCELEASLCSDDCL 120  
Db 61 ASFPLENCRCDVACKDRCDCDFEDTCVETRIWNCNFRGCELEASLCSDDCL 120  
Qy 121 QKDCADYKSVCGGETSWLEENCDTAAQSCQCPGFDLPVILFSDMGFRAEYLYTWDTL 180  
Db 121 QKDCADYKSVCGGETSWLEENCDTAAQSCQCPGFDLPVILFSDMGFRAEYLYTWDTL 180  
Qy 181 MENTNKLTGCIHSHKYMAMYPTKTFPHNYITVGLYPESHGIIIDNNYDVLNKNFSL 240  
Db 181 MENTNKLTGCIHSHKYMAMYPTKTFPHNYITVGLYPESHGIIIDNNYDVLNKNFSL 240  
Qy 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPFSIYMPYNGSVPFERIS 300  
Db 241 SKEQNNPAWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPFSIYMPYNGSVPFERIS 300  
Qy 301 TLLKWLDPKAEPRFYTWYFEEBDSGHAGGPPVSARVICALQVVDHAFGLMEGLKORN 360  
Db 301 TLLKWLDPKAEPRFYTWYFEEBDSGHAGGPPVSARVICALQVVDHAFGLMEGLKORN 360  
Qy 361 LHCNVIILLADHGMDQYCNKMEYTDYFPRNFYFMYEGPAPRIAHNIPHDFFSFNS 420  
Db 361 LHCNVIILLADHGMDQYCNKMEYTDYFPRNFYFMYEGPAPRIAHNIPHDFFSFNS 420  
Qy 421 EEIVANLSCKRPDQHFKEPFLPDLPRHLHYAKNVRIDKVLHFDVQOMLAVRSKNTNCGG 480  
Db 421 EEIVANLSCKRPDQHFKEPFLPDLPRHLHYAKNVRIDKVLHFDVQOMLAVRSKNTNCGG 480  
Qy 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVPFENIEVYNLMCDLRIQAPNNGTHGSLN 540  
Db 481 GNHGYNNEFRSMEALFLAHGSPFKEKTEVPFENIEVYNLMCDLRIQAPNNGTHGSLN 540  
Qy 541 HLLKVPFYPGSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVNMQLNLTQEEI 600  
Db 541 HLLKVPFYPGSHAEVSKFSVCGFANPLPTESLDCFCPLHQNSTQLEQVNMQLNLTQEEI 600  
Qy 601 TATVKVNLPGEPVILQKVDHCLLYHREYVSGFKAMMPWSSYTVPOLGDTSPLPPT 660  
Db 601 TATVKVNLPGEPVILQKVDHCLLYHREYVSGFKAMMPWSSYTVPOLGDTSPLPPT 660

Qy 661 VPDCLRADRVPPPSQKSCSYLADKNITHGFLYPPASNRRTSDQYDALITSNLVPMYEE 720  
Db 661 VPDCLRADRVPPPSQKSCSYLADKNITHGFLYPPASNRRTSDQYDALITSNLVPMYEE 720  
Qy 721 FRKWDYFHSVLLIKHATERGNVVSQPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
Db 721 FRKWDYFHSVLLIKHATERGNVVSQPIFDYNDGHFDAPDEITKHLANTDVPITPHY 780  
Qy 781 FVVLTSCKNKSHTPENCPCGMLDVLFPFIIPHRPTNVSCPEGKPEALWVEERFTHIAHVR 840  
Db 781 FVVLTSCKNKSHTPENCPCGMLDVLFPFIIPHRPTNVSCPEGKPEALWVEERFTHIAHVR 840  
Qy 841 DVELLTGDLDFYQDKVQPVSEILQKTYLPTPETTI 875  
Db 841 DVELLTGDLDFYQDKVQPVSEILQKTYLPTPETTI 875

Search completed: July 6, 2004, 13:26:07  
Job time : 59 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:11:52 ; Search time 26 Seconds

(without alignments)  
3237.215 Million cell updates/sec

Title: US-10-005-480A-743

Perfect score: 4804  
Sequence: 1 MESTLTATEOPVKNTLKK.....QPVSEILQILKLYLPPTFFETI 875

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2833366

Minimum DB seq length: 0

Maximum DB seq	length: 0
Maximum DB seq	length: 200000000

**SECRET**

Post-processing: Minimum Match 0%

Maximum Match 100%

+  
 C  
 E  
 C  
 +  
 C

Database : PIR\_78:★

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

2:  $p_{1r2}^*$ :  
3:  $p_{1r3}^*$ :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3972	82.7	875	1	A57080	cell surface antigen
2	2498.5	52.0	925	1	A39216	nucleotide diphosph
3	2411.5	50.2	905	1	A27410	nucleotide diphosph
4	2043.5	42.5	885	1	A55453	nucleotide diphosph
5	2040.5	42.5	915	1	A55144	autotaxin precursor
6	743.5	15.5	496	2	T09931	probable phosphodi
7	737.5	15.4	479	2	T03293	probable phosphodi
8	712.5	14.8	457	2	T09932	probable phosphodi
9	680.5	14.2	461	2	T09933	probable phosphodi
10	644	13.4	477	1	A39390	probable phosphodi
11	637	13.3	485	2	T40657	probable phosphodi
12	632.5	13.2	477	1	A59391	probable phosphodi
13	632	13.2	300	2	A41179	protein kinase PC
14	630.5	13.1	429	2	T33724	probable phosphodi
15	629.5	13.1	453	1	A59389	probable phosphodi
16	623	13.0	829	2	T19494	hypothetical prote
17	553	11.5	451	2	F87407	probable phosphodi
18	508	10.6	674	2	T19495	hypothetical prote
19	499.5	10.4	433	2	B82537	probable phosphodi
20	434	9.0	743	2	S19437	hypothetical prote
21	412.5	8.6	493	2	S50443	probable phosphodi
22	374	7.8	614	2	T30973	hypothetical prote
23	301	6.3	96	2	A25274	phosphodiesterase
24	222.5	4.6	434	2	F69558	ap superfamily lim
25	222.5	4.6	453	2	T16795	hypothetical prote
26	212	4.4	133	2	T09934	hypothetical prote
27	209	4.4	422	2	A81252	pyrophosphatase ho
28	195.5	4.1	422	2	AH1614	weakly pyrophospha
29	150	3.1	360	2	T20867	hypothetical prote

30	146.5	3.0	370	2	C71052
31	138.5	2.9	369	2	A75091
32	135	2.8	476	2	JC7189
33	132	2.7	474	2	A57480
34	129	2.7	369	2	A34614
35	126	2.6	2004	2	AC0314
36	121	2.5	4543	1	A53102
37	120.5	2.5	478	2	C75099
38	118	2.5	781	2	S43534
39	117	2.4	696	2	T27402
40	116.5	2.4	758	2	S62432
41	116	2.4	704	2	A34287
42	116	2.4	2150	2	T32497
43	115.5	2.4	454	2	A46498
44	115	2.4	397	2	T35880
45	114.5	2.4	1327	2	T14594
46	114	2.4	1327	2	T14594
47	113.5	2.4	1327	2	T14594
48	113	2.4	1327	2	T14594
49	112.5	2.4	1327	2	T14594
50	112	2.4	1327	2	T14594
51	111.5	2.4	1327	2	T14594
52	111	2.4	1327	2	T14594
53	110.5	2.4	1327	2	T14594
54	110	2.4	1327	2	T14594
55	109.5	2.4	1327	2	T14594
56	109	2.4	1327	2	T14594
57	108.5	2.4	1327	2	T14594
58	108	2.4	1327	2	T14594
59	107.5	2.4	1327	2	T14594
60	107	2.4	1327	2	T14594
61	106.5	2.4	1327	2	T14594
62	106	2.4	1327	2	T14594
63	105.5	2.4	1327	2	T14594
64	105	2.4	1327	2	T14594
65	104.5	2.4	1327	2	T14594
66	104	2.4	1327	2	T14594
67	103.5	2.4	1327	2	T14594
68	103	2.4	1327	2	T14594
69	102.5	2.4	1327	2	T14594
70	102	2.4	1327	2	T14594
71	101.5	2.4	1327	2	T14594
72	101	2.4	1327	2	T14594
73	100.5	2.4	1327	2	T14594
74	100	2.4	1327	2	T14594
75	99.5	2.4	1327	2	T14594
76	99	2.4	1327	2	T14594
77	98.5	2.4	1327	2	T14594
78	98	2.4	1327	2	T14594
79	97.5	2.4	1327	2	T14594
80	97	2.4	1327	2	T14594
81	96.5	2.4	1327	2	T14594
82	96	2.4	1327	2	T14594
83	95.5	2.4	1327	2	T14594
84	95	2.4	1327	2	T14594
85	94.5	2.4	1327	2	T14594
86	94	2.4	1327	2	T14594
87	93.5	2.4	1327	2	T14594
88	93	2.4	1327	2	T14594
89	92.5	2.4	1327	2	T14594
90	92	2.4	1327	2	T14594
91	91.5	2.4	1327	2	T14594
92	91	2.4	1327	2	T14594
93	90.5	2.4	1327	2	T14594
94	90	2.4	1327	2	T14594
95	89.5	2.4	1327	2	T14594
96	89	2.4	1327	2	T14594
97	88.5	2.4	1327	2	T14594
98	88	2.4	1327	2	T14594
99	87.5	2.4	1327	2	T14594
100	87	2.4	1327	2	T14594
101	86.5	2.4	1327	2	T14594
102	86	2.4	1327	2	T14594
103	85.5	2.4	1327	2	T14594
104	85	2.4	1327	2	T14594

## ALIGNMENTS

## RESULT 1

A57080 cell surface antigen RB13-6 - rat  
N;Contains: phosphodiesterase I (EC 3.1.4.1)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999  
C;Accession: A57080  
R;Deissler, K.; Lottspeich, F.; Rajewsky, M.F.  
J. Biol. Chem. 270, 9849-9855, 1995  
A;Title: Affinity purification and cDNA cloning of rat neural differentiation and tumor  
A;Reference number: A57080; MUID:95247775; PMID:7730366  
A;Accession: A57080  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-875 <DE1>  
A;Cross-references: GB:247987; NID:G806378; PIDN:CAA88029.1; PID:G806379  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: exonuclease; glycoprotein; phosphoprotein; phosphoric diester hydrolase; su  
F;1-22/Domain: cytosolic #status predicted <CVT>  
F;23-45/Domain: transmembrane #status predicted <TM>  
F;46-875/Domain: extracellular #status predicted <EXT>  
F;51-94/Domain: somatomedin B homology <SBH>  
F;95-138/Domain: somatomedin B homology <SBH2>  
F;206/Binding site: AMP (Thr) (covalent) #status predicted  
F;937/Binding site: AMP (Thr) (covalent) #status predicted  
F;937/280, 289, 533, 574, 594, 702, 789/Binding site: carboxhydrate (Asp) (covalent) #status predicted

Query Match 82.7%; Score 3972; DB 1; Length 875;  
Best Local Similarity 81.2%; Pred No. 7.2e-272;  
Matches 711; Conservative 77; Mismatches 86; Indels 2; Gaps 2;

QY	1	MESTTTLTATEQPVKNTLKKYKACIVLLALLVIMSLGLGLGLURKLBKQ-GSCRKKCF	59
DB	1	MDSLALATEEPIKDSLKKYKILCAVLLALLVIMSLGLGLURKLPBHGTSCKRKKCF	60
QY	60	DASFRGLNCRCDVACKDRGDCDCCWDEDTCVSESTRIMCNKFRCGEETRLAEALSCSDCC	119
DB	61	DSSHRGLEGCDCSGCTDRGDCDCCWDEDTCVKSTQWTNCNFRCGEETRLAEALSCSADCC	120
QY	120	LQKDKCCADYKSVCOGSETSMLENCNPTAOSQCPGFDLPVTLFMSDGFRAEYLVWTDT	179
DB	121	LQKDKCCTDYKAVCOGEVPMVTEACASSQBQCPGFDQPPVTLFMSDGFRAEYLVQWST	180
QY	180	LMPNINKLKTGCIHSKWYRAMYPTKFPNNHYITVGLYPESHGIIIDNNYDVNLKNFSL	239
DB	181	LLPNINKLKTGLGHSKWYRAMYPTKFPNNHYITVGLYPESHGIIIDNNYDVNLKNFSL	240
QY	240	SKSQNNPAMWHGQPMWLITAMYQGLKAATYFWPGSEVAINGPSPTIYMPYNGSVFPBRI	299
DB	241	SSVEKSNPAMWSGGPIWLTAMYQGLKAASYWMLGSVAVNGSPFNTRYNSVSPYESRI	300
QY	300	STLLKWLNDLPAKERPRFYTWYFBEPDSSGHAGSPVARIKALQVVDHAFGLMELGLQR	359



Db 828 PTHFFIVLTSGKDTSTQTLPHCN-LDTLAFILPHRTDENSECVRHKDHSSWVEBLLMLHR 886

Qy 837 ARVRDVELLTGLDFYQDKVQVSEILQLKTLPTF 871

Db 887 ARITDVEHITGLSFYQQRKEPVSIDLKTLPTF 921

RESULT 3

A27410

nucleotide diphosphatase (EC 3.6.1.9) - mouse

N:Alternate names: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.1.1)

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002

C:Accession: A27410; 159055; S38354

R:van Driel, I.R.; Goding, J.W.

J. Biol. Chem. 262, 4882-4887, 1987

A:Title: Plasma cell membrane glycoprotein PC-1. Primary structure deduced from cDNA cloc

A:Reference number: A27410; MUID:87165906; PMID:3104326

A:Accession: A27410

A:Molecule type: mRNA

A:Residues: 1-905 <VAN>

A:Cross-references: GB:J02700; NID:G200236; PIDN:AAA39893.1; PID:G200237

A:Note: The authors translated the codon CAG for residue 24 as Glu

R:van Driel, I.R.; Wilks, A.F.; Pieterz, G.A.; Goding, J.W.

Proc. Natl. Acad. Sci. U.S.A. 82, 8619-8623, 1985

A:Title: Murine plasma cell membrane antigen PC-1: Molecular cloning of cDNA and analysis

A:Reference number: 159055; MUID:86094275; PMID:3001713

A:Accession: 159055

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 203-219 <RES>

A:Cross-references: GB:M12552; NID:G200234; PIDN:AAA39892.1; PID:G200235

R:Bell, S.I.; van Driel, I.R.; Goding, J.W.

Eur. J. Biochem. 217, 421-428, 1993

A:Title: Identification and characterization of a soluble form of the p.lasma cell membra

A:Reference number: S38354; MUID:94039066; PMID:8223581

A:Accession: S38354

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 35-219 <BEL>

A:Cross-references: EMBL:104516

C:Genetics:

A:Introns: 62/3; 87/1; 126/1; 168/1; 188/2

C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology

C:Keywords: coenzyme A; Glycoprotein; phosphoprotein; phosphoric diester hydrolase; tran

F:127-170/Domain: somatomedin B homology <SBH1>

F:161,267,323,459,567,624/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:238/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 50.2%; Score 2411.5; DB 1; Length 905;

Best Local Similarity 50.7%; Pred. No. 7.9e-162;

Matches 445; Conservative 152; Mismatches 251; Indels 29; Gaps 10;

Qy 10 EOPVKK-----NTLKKYKIAICIVLLALVIMSLGLGLGLGLKLEKQSGCRKKCF 59

Db 39 EEPLEKADGARPADKNT---YKVLISLVSVCLTILGICFLGKPSCAKEVKSGKGRCF 95

Qy 60 DASFRGLNCRCDVACKDRGDCWDPEFICVETRIWMCKNRCFCGETRLEASLCSGDDC 119

Db 96 ERTF---SNCRDAACVSLGNCCLDFQETCCEPETHWTNCKFKCGEKRLSRFVSCADDC 152

Qy 120 LQKDCADYKVCQGETSLMELENCDTAQOSQCEGFDPPLPVILFMDGFRAYLYTWDT 179

Db 153 KTHNDCCINYSVCQDKSKSVEITCSIDTPECPAFESPPTLLFSLDGFRAFLHTWG 212

Qy 180 LMPNINKLTCGTHSKYKRMAYPTKTPPNHYITVTGLYPESHGIIINNYDVNLKNVSL 239

Db 213 LLPVSKLNCGTGTYTKMRPMYPTKTPPNHYISVTGLYPESHGIDNRMYPDKMNASFSL 272

Qy 240 SSKQNNPAMWHPWMLTAWYGLKAATYFPGSEVATNGSPESYMPYNGSVPPPEERI 299

Db 273 KSKEFNFLWKQPIWLTANHQEVKSGTYFWGSDVEIDGILPDIYKYNGSVPPPEERI 332

Qy 300 STILLKMLDLPKARPRFYTWYFEEPPSSGHAGGPGVSGARVVKALQVVDHAFGLMMEGLKOR 359

Db 333 LAVLEMLQLPSHERPHFYTLYLEEPDSSGHSGFVSSEVIKALQKQVDRIVGLVLMMDGLKOL 392

Qy 360 NLHNCVNIILLADHGMDOQTYCNKMEYMTDYFPRINFYMYVEGAPRIRAHNIPHDFFSFN 419

Db 393 GLDKCLNLLISDHGMEQSCKYVLYNKYLDGVNNVYVYGAARLPDVPETYYSFN 452

Qy 420 SEBIVNLSCRKPDQHFYPYLPDLPKRLHYAKNVRIDKVLHLEVDQOM-LAVRSKNTNC 478

Db 453 YEALAKNLSCREPNQHFRPYLPKFLPHFAKSDRIEPLTFYLDLPQWQLALNPSEKRYC 512

Qy 479 GCGNHGVNNEFRSMEALFLAHGSPKTEKTEVEPENTVYVNLMDLRLQCPAPNNGTHGS 538

Db 513 GSGFHGSDNLFNSMQALFIGYGAFAKHGAEDVSFEMIEVNLMDLRLGLLIPAPNNGSHGS 572

Qy 539 LNHLKVPFPEPESHABEVSKFSCGFANPLPTEBSLDCFCFHLQNSQLQEOVNQMLNLTOE 598

Db 573 LNHLKKPIYNPSHPKKEGFLSOC---PIKSTNDLGCCTCDPWIVPIKDFEQLNLTE 628

Qy 599 EITATVKVNLPPGRPRVQLKQNDHCLLYHREYVSGFKAMRMPWSSYITVPQLGDTSPLP 658

Db 629 D-DDIYHNTVPYGGPRILLKQHEVCLLQQQOFLTGYSLDLMLPLWASYTF--LRNQOFSR 685

Qy 659 PTVDPCLRADRVPPPSQKCSFVLADKNITHGFLYPPASNRITSDSOY-DALITSLNVP 717

Db 686 DDFSNCMYQDLRPLSPVHKCSYKSKLSYGFLTPPRLNRYSNHITSEALLTSNVP 745

Qy 718 YEEFRKMDYFHSVLLIKHATERNGVNVSGPIFDYNYDGHFADPDEITHK---LANTDV 774

Db 746 YQSPQVIMWHLDTLLQRYAHNRNGINNVSGPFDYDQRYDSLEILKQNSVIRSQEI 805

Qy 775 PIPHYFVLTSCNCKSHTPENGMDVLPEIIPHPTNVESCEPKPALVVEERFTA 834

Db 806 LIPHFVITVTSKQLSETPLECA-LESSAYILPHRPDNIESCTHGKRESSWVEELLTL 864

Qy 835 HIARVRDVELLTGLDFYQDKVQVSEILQLKTLPTF 871

Db 865 HRAVRTDVELITGLSFYQDRQESVSELRLKTLPIF 901

## RESULT 4

A55453

nucleotide diphosphatase (EC 3.6.1.9) - rat

N:Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002

C:Accession: A55453; J00187

R:Narita, M.; Goji, J.; Nakamura, H.; Sano, K.

J. Biol. Chem. 269, 28235-28242, 1994

A:Title: Molecular cloning, expression, and localization of a brain-specific phosphodie

A:Reference number: A55453; MUID:95050605; PMID:7961762

A:Accession: A55453

A:Molecule type: mRNA

A:Residues: 1-885 <NAR>

A:Cross-references: GB:D28560; NID:G464196; PIDN:BAA05910.1; PID:G464197

R:Narita, M.; Goji, J.; Sano, K.; Nakamura, H.

submitted to JPIB, February 1994

A:Description: Cloning and expression of brain-specific nucleotide diphosphohydrolase.

A:Reference number: J00187

A:Accession: J00187

A:Molecule type: mRNA

A:Residues: 1-66,'Q',68-81,'T',83-94,'C',96,'A',98-195,'A',197-514,'E',516-621,'E',623-6

A:Experimental source: strain Sprague-Dawley

C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology

C:Keywords: coenzyme A; exonuclease; glycoprotein; phosphoprotein; phosphoric diester hy

F:54-97/Domain: somatomedin B homology <SBH1>

F:98-141/Domain: somatomedin B homology <SBH2>

F:53,150,396,408,522,608,829/Binding site: carbohydrate (Asn) (covalent) #status predic

F:207/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 42.5%; Score 2043.5; DB 1; Length 885;



482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPEENIEVNLMDLLRIOPAPNNGTHGSLNH 541  
 526 DHGDFNKVNSMOTVFGVGPFTFKYKTKVPPFENIELYNVMDLLGLKLPAPNNGTHGSLNH 585  
 542 LKVPFYEFSHAIEVSFKFSVOCFANPLPTESLDQFC- PHLQNSTQLEQVOMLNLTOEEI 600  
 566 LLRTWTPRTPEEVTRENYRPGIMYLQSDFDLCTCDDKVEPKNKLDELNKLHTK---- 641  
 601 PAVKVNLPFGPRVQLKQVHCLLYHREYSGFGKAMRPMWSSYTYPOLGDTSPPLPT 660  
 642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFSGYSEIFLMLTWTSYTSKQAEVSSVDPH 700  
 661 VPDCLRADVRVPPSSQKCSFYLDADKNITHGLYPFASNR\*SDSQYDALITSNLVPMYEE 720  
 701 LTVSCVRPDRVSPFSQNLAYKNDQKMSYGLFPYLLSSPEAKYDAFLVTVNVMYMPA 760  
 721 FRKMDYFHSVLLIKHATERNGVNVSGPIFYNDYDGHADDEITKHLANDYDPIPHY 780  
 761 FRKVMNYFORVLVKKYASERNVWISGPIFYDYDGLHDTEDKIKQYVEGSSIPVPHY 820  
 781 FVVLTSCKNSHTPENCPCWGLDVLFPILPHRTNVESCPGKPEALWVEERFTHAIAVR 840  
 821 YSIITSLDFTOPAKCDGPLSVSSFILPHRPDNESCNSSDESKWVEELMKMTARVR 880  
 841 DVELLTGLDFYCDKQVQVSEIILQKTYLTFPTTI 875  
 881 DIEHLTSLDFFPKTSRSYSEIILTKYLTHTYSEI 915

RESULT 6  
 T09931  
 probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Aug-2002  
 C:Accession: T09931  
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
 submitted to the Protein Sequence Database, June 1999  
 A:Reference number: Z16897  
 A:Accession: T09931  
 A:Molecule type: DNA  
 A:Residues: 1-496 <REV>  
 A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.190  
 A:Experimental source: cultivar Columbia; BAC clone T16L4  
 C:Genetics:  
 A:Gene: ATSP:T16L4.190  
 A:Map position: 4  
 C:Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4  
 C:Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 15.5%; Score 743.5; DB 2; Length 496;  
 Best Local Similarity 41.0%; Pred. No. 1.4e-44;  
 Matches 159; Conservative 71; Mismatches 141; Indels 17; Gaps 9;

QY 157 DLPPVILFSDGPRAEYLYTWDTLMPNINKLTCGHSKY-MRAMPYTKTFPNHYTIVTG 215  
 DB 105 DKPVLLISDGF--FGYQFKYLPISHLRIANGTEAETGLIPVPTLTLPNHYSIIVG 162  
 QY 216 LYPESHGIIIDNNMYDNLNKNFSLSSKEQNPAWHGQPMWLTAMQGLKAAATYFWPGSE 275  
 DB 163 LYPAYHGIINHHFVDPETGNVFTMASHE---PEWNLGEPLWETVWQGLKAAATYFWGSE 219  
 QY 276 VAINGSF---PSIYMPYNGSVPEERISTLLKWLDPKAEPRFYTYMYPEEPSSGHAGG 332  
 DB 220 VH-KGSWNCPOGLQNTYNGSVFFDDRVDTLTLSYFDLPSNIPSPMTLYFEDPDHQHQVG 278  
 QY 333 PVSAVRIKALQVVDHAFQMLMEGLKQRLNHCNVNIIADHGNDQTYCNKMEYMTDYFPR 392  
 DB 279 PDDQITEAVVNIIDRLIGRLIDGLEKRGIPEDVTMIVGDHGMVGTCDKLLVLLDLAPW 338  
 QY 393 INF---FMYEGGPAPRIAHNIPHDFFSNSEIEIVNLSCKPD--QHFKPYLTPDLPKR 447  
 DB 339 IKIPSSWQYTYTLLAIQPPS-GHDAADIVA-KINEGLSGKVENGYLKYLVKLEDLPKR 396

448 LHYAKNVRIDKVLHFDVQOQWLAVRSKSNNTNCGGNHGYNNFRSMEAIFLAHGPSFKEKT 507  
 397 LHYVDSRIPPIIGLVEGFKVEQKSKAKCEGAGHYDNAAFSSMTITFIGHGPMFSKGR 456  
 508 EYEPFENIEVNLMDLLRIOPAPNNGT 535  
 457 KVPSEFNVQIYNVISSILGLKAAPNNGS 484

RESULT 7  
 T03293  
 probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) - rice  
 C:Species: Oryza sativa (rice)  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Aug-2002  
 C:Accession: T03293  
 R:Hsing, Y.C.; Teao, C.V.; Chow, T.; Hsieh, J.; Chen, Z.  
 submitted to the EMBL Data Library, April 1995  
 A:Description: Rice early embryogenesis gene.  
 A:Reference number: Z14889  
 A:Accession: T03293  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-479 <HSI>  
 A:Cross-references: EMBL:U25430; NID:g818848; PID:g818849  
 A:Experimental source: strain Tainung 67  
 C:Genetics:  
 A:Gene: OS84  
 C:Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4  
 C:Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 15.4%; Score 737.5; DB 2; Length 479;  
 Best Local Similarity 42.0%; Pred. No. 3.4e-44;  
 Matches 163; Conservative 64; Mismatches 140; Indels 21; Gaps 10;

QY 159 PPVILFSDGPRAEYLYTWDTLMPNINKLTCGHSKY-MRAMPYTKTFPNHYTIVTGLY 217  
 DB 90 PVVILISDGFPGYQKAAAT--PHIHLIGNGTSAAATGLVPIPTLTLPNHYSIATGLY 147  
 QY 218 PESHGIIIDNNMYDNLNKNFSLSSKEQNPAWHGQPMWLTAMQGLKAAATYFWPGSEVA 277  
 DB 148 PSSHGIIINNYFPDPIGDIYFTNSSHE---PKWNLGEPLWVTAADQGLKAAATYFWPGSEVK 204  
 QY 278 INGSF--PSIY-MPYNGSVPEERISTLLKWLDPKAEPRFYTYMYPEEPSSGHAGGPV 334  
 DB 205 -KGSWDCPKYCRHYNGSVPEERVDAILGVDFLPSDEMPQFLTYFEDPDHQHQVGPDP 263  
 QY 335 SARVRIKALQVVDHAFQMLMEGLKQRLNHCNVNIIADHGNDQTYCNKMEYMTDYFPRIN 394  
 DB 264 DPAITEAVVRIIDEMIGRLIAGEGVFEDVNVILVGDHGMVGTCDKLLVFLDELAPWIK 323  
 QY 395 F---FMYEGGPAPRIAHNIPHDFFSNF--EEIVNLSCKPD--QHFKPYLTPDLPKR 447  
 DB 324 LEEDVLSMTPLAIR----PPDDMSLPDVVAKNNEGLSGKVENGEYLRWYKLEDLPKR 379  
 QY 448 LHYAKNVRIDKVLHFDVQOQWLAVRSKSNNTNCGGNHGYNNFRSMEAIFLAHGPSFKEKT 507  
 DB 380 LHYADSYRIPPIIGLVEEGYKVEKRSKNECGAGHYDNAAFSSMTITFIAHGPFEGR 439  
 QY 508 EYEPFENIEVNLMDLLRIOPAPNNGT 535  
 DB 440 VVPSFENVEIYNVIASILNLEPAPNNGS 467

RESULT 8  
 T09932  
 probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Aug-2002  
 C:Accession: T09932  
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
 submitted to the Protein Sequence Database, June 1999  
 A:Reference number: Z16897  
 A:Accession: T09932





Db 299 EIPERWHYKNDRVQIVAVADEGWYILQKSD-DFLGNHGYDNALAEHPIFLAHGP 357  
QY 502 SFKEKEVEPFENIEVYINLMCDLRIQAPNNGTHGSLNHLKLVFPFEPASHAEVSKFSV 561  
Db 358 AFKFNFTKEAMSTDLISLCHLNLTALPHNGSFNNVQDLL----- 399  
QY 562 CGFANPLPESLDCFCFHLQNSTOL 586  
Db 400 -SSATPKP-----IPYTSQSTLL 416

## RESULT 11

T40657  
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) SPBC72  
C/Species: Schistosoma japonicum  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 16-Aug-2002  
C/Accession: T40657  
R/Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.  
submitted to the EMBL Data Library, November 1998  
A/Reference number: Z21889  
A/Accession: T40657  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Cross-references: EMBL:AL034352; PIDN:CAA22177.1; GSPDB:GN00067; SPDB:SPBC725.05c  
A/Experimental source: strain 972h-; cosmid c725  
C/Genetics:  
A/Gene: SPDB:SPBC725.05c  
A/Map position: 2  
C/Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4  
C/Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 13.3%; Score 637; DB 2; Length 485;

Best Local Similarity 37.5%; Pred. No. 4.2e-37; Matches 146; Conservative 64; Mismatches 165; Indels 14; Gaps 8;

QY 161 VILFSDMGFRAEYLYTWTDLMPNINKLTCGHSKYMRAMYPTKTPPNHYTVITGLYPS 220  
Db 76 VIVISLDGFADLYRGFT--PNLLSLAENNVHVPFLIPSPFSITFPNHYTVITGLYPS 133  
QY 221 HGIIDNMVDNINKNFSLSSKEQN-NDPAW-HGQPMWLTAMYQGLKAATYWPQSEVAL 278  
Db 134 HGVSNFFDPVTGKQFVMSPECNDPTWMDXGEPIWNAERNVRSVHVMWPGNEVEN 193  
QY 279 NGSPFSTYMPYNGSVFPFEERISILLKWLDPKAERPRFTVWTFEEDSSCHAGGPVSARV 338  
Db 194 HGRPTYSDGPNFTILREKORILEWLDLPKDRFQLLAVAPHVDMVGHAFGDPSPFL 253  
QY 339 IKALQVVDHAFGLMEGLKQRLNHCNVIILLADHGMDQTYCNKMEYMTDYPFRINFFYM 398  
Db 254 NIITQEVDIIVIGELIEGLKRNKDKHVIIFLSDHGWAPTSNRLIWLDMNLSAVAH 313  
QY 399 YEGPAPRIRAHNIPHDFFSFNSEIYVNLSCRKPD-QHKPYLTPDLPKRLHYAKNVRID 457  
Db 314 DAMPLQGFROGESDLDDEYIY--ESLVNYSRSSLPSEAENNNVYSKKDIPSRWHYNNHRIA 371  
QY 458 KVLHFLVDQQLAV-----RSKSNNTCCGGNHGYNNEFRSMERAIPLAHGPFK--EKTEVEP 511  
Db 372 PVWMPIDVGSLSVMDLHSDPELVEPLGVHGYDNLSPVWRALFIASSGSFKNPKGKLAP 431  
QY 512 FENIEVYINLMCDLRIQAPNNGTH-GSL 539  
Db 432 FONTEIYGILSHILDLPAPQPNNGTYEGAL 460

## RESULT 12

A59391  
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) 5 - hu  
N/Alternate names: ectonucleotide pyrophosphatase/phosphodiesterase 5; NPP5  
C/Species: Homo sapiens (man)  
C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 16-Aug-2002  
C/Accession: A59391  
R/Shalley, C.

submitted to the EMBL Data Library, September 1999  
A/Description: Contains actin, gamma pseudogene 9, a gene similar to PLASMA-CELL MEMBRAN  
A/Reference number: A59391  
A/Accession: A59391  
A/Molecule type: DNA  
A/Residues: 1-477 <SNA>  
A/Cross-references: GB:AL035701; MID:G5924007; PIDN:CAE56566.1  
R/Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M.  
J. Biol. Chem. 276, 1361-1368, 2001  
A/Title: Structural and catalytic similarities between nucleotide pyrophosphatases/phosph  
A/Reference number: A59390; MUID:21125673; PMID:11027689  
A/Contents: annotation  
C/Genetics:  
A/Gene: ENPP5  
A/Map position: 6p11.2-6p21.1  
C/Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4  
C/Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase  
F/72/Active site: Thr #status predicted

Query Match 13.2%; Score 632.5; DB 1; Length 477;

Best Local Similarity 34.4%; Pred. No. 8.6e-34; Matches 140; Conservative 80; Mismatches 137; Indels 33; Gaps 10;

QY 156 FDLPP-----VILFSDMGFRAEYLYTWTDLMPNINKLTCGHSKYMRAMYPTKTPPNHYT 211  
Db 21 FSLQPDQKQVLLVSFGDFRWDLYKVPT--PAPHYIMKYGVHVQVTNVFTKTYPNHYT 78  
QY 212 IVTGLYPESHGIIIDNMVDNINKNFSLSSKE-QNNPAWVHGQPMWLTAMYQGLKAATYF 270  
Db 79 LVTGLFAENHGVIVANDMPDIRNKSFLSDHNVNIVDSKFWEATPIWITNQAGHTSGAAM 138  
QY 271 WPGSEVAINSGFPIYMPYNGSVFPFEERISILLKWLDPKAERPRFTVWTFEEDSSGHA 330  
Db 139 WPGTDVILKHFPHYMPYNSVSFEDRVAKIEM--FTSKEPINLGLLYWEDPDDMGHH 196  
QY 331 GGPVSARVIAKALQVVDHAFGLMEGLKQRLNHCNVIILLADHGMDQTYCNKMEYMTDIF 390  
Db 197 LGPDSPLMGPIVSDIDKGLYLIQMLKAKLWNTLNIITSDHGMTQCSERLIELDQYL 256  
QY 331 PRINFFYVMEGPAPRIRAHNIPHDFFSFNSEIYVNLSCRKPDQHKPYLT---PDLPK 446  
Db 257 DRDHTLIDQSPVAI-----LPKE---GKFDEVYEALT-----HAPNLTIVYKKEDEVPE 303  
QY 447 RLHYAKNVRIDKVLHFLVDQQLAVRSKSNNTCCGGNHGYNNEFRSMERAIPLAHGPFKEK 506  
Db 304 RWHYKNSRIQPIITAVADEGWHILQKSD-DFLGNHGYDNALADHPIFLAHGPAFRKN 362  
QY 507 TEVEPFENIEVYINLMCDLRIQAPNNGTHGSLNHLK-----VPF 547  
Db 363 FSKAAMNSTDLYPLCHLLNITAMPHNGSFNNVQDLLNSAMPVVPY 409

## RESULT 13

A41179  
protein kinase PC-1 (EC 2.7.1.1) - bovine (fragments)  
N/Alternate names: MAPK; major acidic fibroblast growth factor-stimulated phosphoprotein  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 28-May-1992 #sequence\_revision 22-Apr-1995 #text\_change 30-Apr-1999  
C/Accession: A41179; A49308  
R/Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 266, 16791-16795, 1991  
A/Title: The plasma cell membrane glycoprotein, PC-1, is a threonine-specific protein k  
A/Reference number: A41179; MUID:91358477; PMID:1715869  
A/Accession: A41179  
A/Molecule type: protein  
A/Residues: 1-26;36-56;59-67;68-133;134-144;145-267;268-300 <ODA>  
A/Experimental source: liver  
R/Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 268, 27318-27326, 1993  
A/Title: The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein fro  
A/Reference number: A49308; MUID:94086550; PMID:7505270  
A/Accession: A49308



A:Molecule type: protein  
A:Residues: 27-35, 'X', 37-58 <OD2>  
A:Experimental source: liver

C>Note: sequence extracted from NCBI backbone (NCBIP:141583)  
C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C:Keywords: glycoprotein; phosphoprotein; phosphotransferase  
F:1-25/Domain: somatomedin B homology (fragment) <SBH>  
F:36/Binding site: AMP (Thr) (covalent) (status predicted)

Query Match 13.2%; Score 632; DB 2; Length 300;  
Best Local Similarity 21.8%; Pred. No. 4.9e-37;  
Matches 174; Conservative 51; Mismatches 71; Indels 504; Gaps 13;

QY 57 KCFDAAGFGLNCRCDVACHDRGDCWDPEDTCVETRIWCMKFCGETRLEASLCSCS 116

Db 2 KCFERTF---GNCRCDAACVGLGNCCLD----- 26

QY 117 DDCLQKDCADYKSVCOGETSLEENCDAQAQSCQCEGFDLPPVILFMDGFRAEYLTY 176

Db 27 ----- 26

QY 177 WDTLMFNINKLTCGHSKYMRYMPTKTPNNHYTIVTGLYPESHGIIIDNNYDVLNKN 236

Db 27 -----NMREYVPTKTPNNHYTIVTGLYPESHGIIIDNK----- 58

QY 237 FSLSSKEQNNPAMWHGQPMWLTAMYQGLKAAATYFPGSEVAINGSPSPSYMPYNGSVRPE 296

Db 59 -----GEPIWLT-----KSGIFFWPGSDVKINGFPDIYXV---SVFPE 95

QY 297 ERITLLKWLDPKAERPRFYMYFPEPDSSGHAGGVSARVAKLVVVDHAFGLMEGL 356

Db 96 ERITAILKWLQLP----- 108

QY 357 KQNLNLCNVIILLADHGMDQTYCNKMYMTDYFFRINFYMEGPAPIRAHNIPHPFF 416

Db 109 KEUNLHRCNLILISDHGMEQSGK-----YY 135

QY 417 SFNSEIIVRLNLSCKRPDQHFKPYLTDPKRLHYAKQVRIDKVLHVPDQW-LAVRSKN 475

Db 136 SPDYEGI-----AKSDIERLTFLDPQWQLNPSER 168

QY 476 TNCGGHGVNNEFRSNEALFLAHGSPFKEKTEVEFENIEVYNLMDLLRIQAPNNGT 535

Db 169 KYCGGFHGSNDLFLNQALFXXXXXXHXSTEVDSFENIEVYNLMDLLNLTAPNNGT 228

QY 536 HGSNLHLKVPFYPERSHAEVSKFVCGFANPLPTESLDCFCPLQNSQLQVQVQMLN 595

Db 229 H----- 229

QY 596 TOEETATVKVNLPGFRVQLQNVHDCLLYHREYVSGFKAMRPMWSSYVTPQLGDS 655

Db 230 ----- 229

QY 656 PUPPTVPDCLRADVRVPPESQKSFYLADKNITHGFLYPPASNTSDSQY-DALITSNL 714

Db 230 -----LSYGFUSPPLHKGSGQVSEALLTNI 257

QY 715 VPMYEEFRKMDYFHSVLLIKHATERNGVNVVSGPIFDYNDGHFADPEITKHLANTDV 774

Db 258 VPMYQSQFQ----- 265

QY 775 PIPTHFVVLTSCKNKSHTPENCQWLDVLPFIIPHRPTNVESCEPKGPEALWVEERTA 834

Db 266 -----VTH-----ESLWVEELKL 279

RESULT 14

T33724

probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) - Zymo

C:Species: Zymomonas mobilis

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 16-Aug-2002

C:Accession: T33724

R:Lee, J.S.; Jin, S.J.; Kang, H.L.; Kang, H.S.

submitted to the EMBL Data Library, August 1998

A:Description: Sequence analysis of 67E10 cosmid clone of Zymomonas mobilis 2M4.

A:Reference number: Z21392

A:Accession: T33724

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-429 <LEE>

A:Cross-references: EMBL:AF086791; NID:G3820581; PID:G3089614; PIDN:AAC70363.1

C:Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4

C:Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 13.1%; Score 630.5; DB 2; Length 429;  
Best Local Similarity 35.4%; Pred. No. 1e-36;  
Matches 135; Conservative 76; Mismatches 143; Indels 27; Gaps 10;

QY 160 PVILFSDGTPRAEVLTYTWTLMFNINKLTCGHSKYMRYMPTKTPNNHYTIVTGLYPE 219

Db 49 PLILISIDGPRADYIKR--GLTFLNLSLAENGSAKYAKMHPSPFSITFPNNHYTIVTGLYPD 106

QY 220 SHGIIDNNMYDVNL--NKNFSLSSKE--QNNPAMW-HGQPMWLTAMYQGLKAAATYFPGSE 275

Db 107 HEGIVGNMDDAHTTPDSHFMSDQQAATDRWNWDEGEPLVTAEGQVVSATMFWPGSE 166

QY 276 VAINGSPSYMPYNGSVVPEERISTLLKWLDPKAERPRFYMYFPEPDSSGHAGGVS 335

Db 167 ADIEGVPEPMWOFDSHVPSERVQVFSWLAYPEKKRPQFILTYPENVDHAGHLYGPD 226

QY 336 ARVICALQVVDHAFGLMEGLKORNLHNCVNIILLADHGMDQTYCNKMYMTDYPPRINF 395

Db 227 QEVNDNLVLDITQIGLVQGLKQKGIK--ANLVISDHGMAATSDRVVALNKILDPGLY 284

QY 396 FYM----YEGPAPIRAHNIPHDFFSNSEIIVRLNLSCKRPDQHFKPYLTDPKRLHYA 451

Db 285 HVTGGAYAGIEPS--SGHSLKDLPLFASHD-----HMQWCPKQIPARFHYG 331

QY 452 KVRIDKVLHVPDQWLAVERSKS--NTNCGGNGHYNNFRSNEALFLAHGSPFKEKTEVE 510

Db 332 QNRPVPAVCAAEVGVSGIMGDDSNAAHATKGNHGYNDQTEMGALFIANGPAQKHVIE 391

QY 511 PFENIEVYNLMDLLRIQAP 531

Db 392 SMDNIDVQLVAQVQLK-AP 411

RESULT 15

A59389

probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) 4 - hu

N:Alternate names: ectonucleotide pyrophosphatase/phosphodiesterase 4; KIAA0879 protein;

C:Species: Homo sapiens (man)

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 16-Aug-2002

C:Accession: A59389; B59391

R:Nagase, T.; Ishikawa, K.; Suyama, M.; Kikuno, R.; Hirose, M.; Miyajima, N.; Tanaka,

DNA Res. 5, 355-364, 1998

A:Title: Prediction of the coding sequences of unidentified human genes. XII. The comple

A:Reference number: A59389; MUID:99156230; PMID:10048485

A:Accession: A59389

A:Molecule type: mRNA

A:Residues: 1-453 <NAG>

A:Cross-references: GB:AB020686; NID:G4240247; PIDN:BAA74902.1

A:Experimental source: adult brain

R:Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M.

J. Biol. Chem. 276, 1361-1368, 2001

A:Title: Structural and catalytic similarities between nucleotide pyrophosphatases/phosph

A:Reference number: A59390; MUID:21125673; PMID:11027689

A:Contents: annotation

R:Smalley, C.

submitted to the EMBL Data Library, September 1999

A:Description: Contains actin, gamma pseudogene 9, a gene similar to PLASMA-CELL MEMBRAN

A:Reference number: A59391

Search completed: July 6, 2004, 13:20:28  
Job time : 29 secs

This Page Blank (uspto)